

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTACAAAAGGTGCAGGTA
TGAGCAGGTCTGAAGACTAACATTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTGG
TGGTTTCAGCAAGGCCTCAGTTCCCTCCTCAGCCCTGTAATTGGACATCTGCTGCTTC
ATATTTCATACATTACTGCAGTAACACTCCACCATAAGACCCGGCTTACCTTATATCAGT
GACACTGGTACAGTAGCTCCAGAAAAATGCTTATTGGGGCAATGCTAAATATTGCGGCAGTT
TTATGCATTGCTACCATTATGTTGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAAC
GTTATCATCAAATTAAACAAGGCTGGCCTGTACTTGAATACTGAGTTGTTAGGACTTTCT
ATTGTGGCAAACCTCCAGAAAACAACCCTTTGCTGCACATGTAAGTGGAGCTGTGCTTACC
TTTGGTATGGGCTCATTATATATGTTGTTCAGACCATCCTTACCAAATGCAGCCAAA
ATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGTTATCTGGTGTGGAGTAAGTGCA
CTTAGCATGCTGACTTGCTCATCAGTTGCACAGTGGCAATTGGACTGATTAGAACAG
AAACTCCATTGGAACCCCAGGGACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAA
TGGTCTATGTCATTTCTTCTTGGTTTCTGACTTACATTGCTGATTTCAGAAAATT
TCTTACGGGTGGAAGCCAATTACATGGATTAACCCCTATGACACTGCACCTGCCCTATT
AACAAATGAAACGAACACGGCTACTTCCAGAGATATTTGATGAAAGGATAAAATATTCTGTAA
TGATTATGATTCTCAGGGATTGGGGAAAGGTTCACAGAAGTTGCTTATTCTCTGAAATT
TCAACCACCTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATAATCAGGAAACATGAAAG
AAGCCATTGATAGATTATTCTAAAGGATATCAAGAAGACTATTAAAAACACCTATGCCT
ATACTTTTATCTCAGAAAATAAGTCAAAAGACTATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFS YITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIA
AVLCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTLFAAHVSGAV
LTFGMGSLYMFVQTILSYQMOPKIHGKQFWIRLLLVIWCGVSALSMLTCSSVLHSGNFTDL
EQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFLTYIRDFQKISLRVEANLHGLTLYDTAPC
PINNERTRLLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGCGGACGCGTGGGGAGAGCCGCAGTCCCCTGCAGCACCTGGGAGAAGGC
 AGACCGTGTGAGGGGCCTGTGGCCCCAGCGTGCTGTGGCTCGGGAGTGGAGTGGAGGC
 AGGAGCCTCCTTACACTCGCC**ATG**AGTTCTCATCGACTCCAGCATCATGATTACCTCCC
 AGATACTATTTGGATTGGCTTCTCATGCGCCAATTGTTAAAGACTATGAGA
 TACGTCAGTATGTTACAGGTGATCTCTCCGTGACGTTGCATTTCTGCACCATGTTG
 AGCTCATCATCTTGAAATCTTAGGAGTATTGAATAGCAGCTCCGTTATTCAGTGGAAAA
 TGAACCTGTGTGAATTCTGCTGATCCTGGTTCATGGGCCTTTACATTGGCTATTTA
 TTGTGAGCAATATCCGACTACTGCATAAACAAACGACTGCTTTCTGTCTTATGGCTGA
 CCTTATGTATTCTCTGGAAACTAGGAGATCCCTCCATTCTCAGCCCCAAACATGGGA
 TCTTATCCATAGAACAGCTCATCAGCCGGTTGGTGTGATTGGAGTGACTCTCATGGCTCTC
 TTTCTGGATTGGTGTCAACTGCCCATACACTACATGTCTTACTCCTCAGGAATGTGA
 CTGACACGGATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGATCATAAGCA
 AAAAGAAAAGGATGGCAATGGCACGGAGAACAAATGTTCCAGAAGGGGGAGTGCATAACAAAC
 CATCAGGTTCTGGGAATGATAAAAGTGTACCACTTCAGCATCAGGAAGTGAAAATCTTA
 CTCTTATTCAACAGGAAGTGGATGCTTGGAGAACATTAAGCAGGCAGCTTTCTGGAAACAG
 CTGATCTATATGCTACCAAGGAGAGAACAAATGAGACTCCAAACCTCAAGGGAAATATTTA
 ATTTCTGGTTACTTTCTCTATTACTGTGTTGGAAAATTTCATGGCTACCCTCAATA
 TTGTTTGATCGAGTTGGAAAACGGATCCTGTCACAAGAGGCATTGAGATCACTGTGAATT
 ATCTGGGAATCCAATTGATGTGAAGTTGGTCCACACATTCCTCATTCTGGAA
 TAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTACCAAGTTCTTATGCCATCT
 CTAGCAGTAAGTCCTCCAATGTCATTGCTCTGCTATTAGCACAGATAATGGCATGTACTTG
 TCTCCTCTGTGCTGATCCGAATGAGTATGCCCTTAGAACACCGCACCATAACTGAAG
 TCCTGGAGAACTGCAGTTCAACTCTATCACCGTTGGTGTGATCTCCTGGTCAGCG
 CTCTCTAGCATACTCTCCTCTATTGGCTCACAAACAGGCACCAGAGAACATGGCAC
 CTT**TGA**ACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAAATTAGATATAAGAGG
 GGGGAAAATGGAACCAGGGCCTGACATTATAAACAAACAAAATGCTATGGTAGCATT
 CACCTTCATAGCATACTCCTCCCCGTCAAGGTGATACTATGACCATGAGTAGCATGCCAGA
 ACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCGTGTGGATATGAGG
 CTGGTAGAGGGCGGAGAGGGAGCCAAGAAACTAAAGGTGAAAAACTACACTGGAACCTGGGGC
 AAGACATGTCTATGGTAGCTGAGCCAACACGTAGGATTCCGTTAAGGTTCACATGGAAA
 AGGTTATAGCTTGCCTTGAGATTGACTCATTAAATCAGAGACTGTAACAAAAAA
 AAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTGGCCGCATGGCCAACT
 TGTTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQLFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEIL
GVLNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWK
LGDPFPILSPKHGILSIEQLISRGGVIGVTLMALLSGFGAVNCPTYMSYFLRNVTDTDILAL
ERRLLQTMDMIISKKRMAMARRTMFQKGEVHNKPSGFWMGIKSVTSASGSENLTLIQQEVD
ALEELSRLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGK
TDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNV
IVLLLQIMGMYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFL
YLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACATAGTT
 CCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCACTGGCCATCTG
 AGGTGTTCCCTGGCTCTGAAGGGTAGGCACG**A****T****G**GCCAGGTGCTTCAGCCTGGTGTGCTT
 CTCACTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTGCGTGAGAAGAGCTTCC
 ATCCAGGTGTATGCAGAATTATGGGGATCACCCCTGTGAGCAAAAAGGCGAACAGCAGCTG
 AATTCACAGAAGCTAAGGAGGCCTGTAGGCTGCTGGACTAAGTTGGCCGGCAAGGACCAA
 GTTGAACAGCCTGAAAGCTAGCTTGAAGACTTGAGCTATGGCTGGGTTGGAGATGGATT
 GTGGTCATCTCTAGGATTAGCCAAACCCAAAGTGTGGAAAAATGGGTGGGTGTCCTGATT
 TGGAAAGGTTCCAGTGAGCCGACAGTTGCAGCCTATTGTTACAACATCTGATACTGGACT
 AACTCGTGCATTCCAGAAATTATCACCAACAAAGATCCCATTCAACACTCAAACGTGCAACA
 CAAACAAACAGAATTATTGTCAAGTGAAGCTGCTGGCATCCCCTACTCTACAATA
 CCTGCCCTACTACTACTCCTCCTGCTCCAGCTTCTATTCCACGGAGAAAAAAATTG
 ATTGTGTCACAGAAGTTTATGGAAACTAGCACCATGTCTACAGAAACTGAACCATTGTT
 GAAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTGGAGGTGCCCCACGGCTCTGCTA
 GTGCTTGCTCTCCTCTTGGTGCAGCTGGTCTGGATTGCTATGTCAAAAGGTAT
 GTGAAGGCCTCCCTTTACAAACAAGAATCAGCAGAAGGAAATGATCGAAACCAAAGTAGTA
 AAGGAGGAGAAGGCCATGATAGCAACCTAATGAGGAATCAAAGAAAATGATAAAAACCA
 GAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCATGCCCTGGAAGCTGAAGTT**T****A****G**ATGAGA
 CAGAAATGAGGAGACACACCTGAGGCTGGTTCTTCATGCTCCTTACCCCTGCCAGCTGGG
 GAAATCAAAGGGCCAAAGAACCAAAGAACAGAAAGAACCTTGGTCTCTAACTGGAATCAGC
 TCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAACATGCCCTCTCTATTGTAACCCT
 GTCTGGATCCTATCCTCCTACCTCAAAGCTTCCCACGGCCTTCTAGCCTGGCTATGTCCTA
 ATAATATCCCACGGAGAAAGGAGTTGCAAAGTGCACGGACCTAAACATCTCATCAGTA
 TCCAGTGGTAAAAGGCCTCCTGGCTGTGAGGCTAGGTGGGTTGAAAGCCAAGGAGTCAC
 GAGACCAAGGCTTCTACTGATTCCGCAGCTCAGACCCCTTCTCAGCTCTGAAAGAGAAA
 CACGTATCCCACCTGACATGTCCTCTGAGCCCGTAAGAGCAAAAGAACATGGCAGAAAAGTT
 AGCCCCTGAAAGCCATGGAGATTCTCATAACTGAGACCTAATCTGTAAAGCTAAAATAAA
 GAAATAGAACAAAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAACACAGACAGG
 GTCAAAGTGTGTTCTGAACACATTGAGTTGGAATCACTGTTAGAACACACACACTTACTT
 TTTCTGGTCTTACCACTGCTGATATTTCTCTAGGAAATATACTTTACAAGTAACAAAAT
 AAAACTCTTATAAATTTCTATTCTGAGTTACAGAAATGATTAAGGAAGATTACT
 CAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATATTCTCCTCAAAAATTGCACATAGTA
 AGTGCCTGTGCAAGGTATTACACTCTGTAATTGAATATTCTCCTCAAAAATTGCACATAGTA
 GAACGCTATCTGGGAAGCTATTTTCTGAGTTGATATTCTAGCTTATCTACTTCCAAACT
 AATTTTATTTGCTGAGACTAATCTTATTCTCTAATATGGCAACCATTATAACCT
 TAATTATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAAGCACATTAA
 AAGTGCCATTAACAAATGTATCACTAGCCCTCCTTCCAACAAGAAGGGACTGAGAGATGC
 AGAAATATTGTGACAAAAATTAAAGCATTAGAAAACCTT

FIGURE 6

MARCFSLVLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQLNFTEAKEACRL
LGLSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNSPKCGKNGVGVLIWKVPVSRQFAA
YCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVSPYSTIPAPTTPPAPA
STSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAA
GLGFCYVKRYVKAFFFTNKNQQKEMIETKVVKEEKANDSNPNEESKTDKNPEESKSPSKTTV
RCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCGCACCGCGGCCACCGCGCCGCTCCGCATCTGCACCCGCAGCCG
 GC GG CCT CCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGAACTCGGTCCAGTCGGG
 GC GG CGG CT GC GGG CGC AGAG CGG AG **ATG** CAG CGG CT TGGG GCC ACC CT GCT GT GC CT GCT GC
 TGG CGG CGG CGG TCCC AC GG CCCC CGC CCC GCT CC GAC GG GAC CT CG GCT CC AGT CA AGC
 CC GG CCC GG CT CT CAG CT ACC CG CAG GAG GAG CC ACC CT CA AT GAG AT GT TCC CG GAG GT TG
 AGG AACT GAT GG AGG AC AC G CAG CAC AA ATT GCG CAG CG CG GT GG AAG AG AAT GG AGG CAG AAG
 AAG CT GCT GCT AA AG CAT CAT CAG AAG TGA AC CT GG CAA ACT TAC CT CC AG CT AT CACA ATG
 AG ACC AAC AC AG AC AC GA AG GT GG AA ATA AT ACC AT CC AT GT GC ACC GAG AA AT TC ACA AGA
 TA ACC AA CA ACC AG ACT GG AC AA AT GG TCT TT CAG AG A CAG TT AT CAC AT CT GT GGG AG AC G
 AAG AAG GC AGA AGG AG CC AC GAG TGC AT CAT CG AC GAG GACT GT GGG CC AG AG GAT GCT CT GCA CCC GGG
 AG TTT GCC AG CTT CAG TAC AC CT GCC AG CC AT GCC GGG CC AG AG GAT GCT CT GCA CCC GGG
 AC AGT GAG TGC TGT GG AG ACC AG CT GT GT CT GGG GT CACT GCA CAA AT GG CC ACC CAG GG
 GC AG CA AT GG GACC AT CT GT GACA ACC AG AG GG ACT GCC AG CC GGG CT GT GCT GT GCT TT CC
 AG AG AGG C CT GCT GT CC CT GT GCA CAC CC CT GCC GT GG AG GG GAG CT TT GCC AT GACC
 CCG CC AG CC GG CT TCT GG AC CT CAT CAC CT GG AG CT AG AG C CT GAT GG AG GCT TGG ACC GAT
 GCC CT GT GCC AGT GG C CT CCT CT GCC AG CCCC AC AG CC AC AG C CT GG TGT AT GT GT GCA AGC
 CG AC CT TC GT GGG GAG CC GT GAC CA AG AT GGG GAG AT CT GCT GCC AG AG AG GT CCC GAT G
 AGT AT GA AG TT GG CAG CTT CAT GG AGG AG GT GCG CC AGG AG CT GG AGG AC CT GG AG AG GAG CC
 TG ACT GA AG AG AT GG CG CT GGG GAG C CT GCG CT GCC GCG CT GCA CT GCT GGG AG GGG AAG
 AG ATT **TAG** AT CT GG ACC AGG CT GT GGG TAG AT GT GCA AT AG AA AT AG CT AATT ATT CCC CA
 GGT GT GT GCT T TAGG CGT GGG CT GACC AGG CT TCT CAC AT CT TCT CC AG TA AG TT CC
 CCT CT GG CT T GAC AG CAT GAG GT GT GCA TT GTC AG CT CCCC AGG CT GT T CT CC AGG C
 TTC AC AGT CT GG TG CT TGG GAG AG TC AGG CAG GG TT AA ACT GCA AGG AG CAG TT GCC ACC C
 GT CC AG AT T ATT GG CT GCT T GC CT TAC CAG TT GG CAG AC AG CC GT T GT T CT AC AT GG CT
 TG AT A ATT GT TT GAG GG GAG GAG AT GG AA AC AAT GT GG AGT CT CC CT TG ATT GG TTT GGG
 AA AT GT GG AG AAG AGT G C C T G C T T G C A A A C A T C A A C C T G G C A A A A T G C A A C A A T G A A T T
 TT CC AC GC AG TT CT T C AT GGG C AT AGG TA AG CT GT G C C T T C AG CT GT T G C A G AT G A A A T G T
 TCT GT T C ACC CT G C AT T AC AT GT GT TT ATT C AT CC AG C AG T GT G C T C AG C T C C T A C C T G T
 G C C A G G G C A G C AT T T C AT AT C C A A G A T C A A T T C C C T C T C A G C A C A G C C T G G G G A G G G G T
 C A T T G T T C T C C T C G T C C AT C A G G G A T C T C A G A G G C T C A G A G A C T G C A A G C T G C T T G C C C A A G T
 C A C A C A G C T A G T G A A G A C C A G A G C A G T T C A T C T G G T T G T G A C T C T A A G C T C A G T G C T C T C
 C A C T A C C C C A C A C C A G C C T T G G T G C C A C C A A A A G T G C T C C C A A A A G G A A G G G A A T G G G A T T
 T T T C T T G A G G C A T G C A C A T C T G G A A T T A A G G T C A A A C T A A T T C T C A C A T C C C T C T A A A A G T A A
 A C T A C T G T T A G G A A C A G C A G T G T C A C A G T G T G G G C A G C C G T C T T C T A A T G A A G A C A A T
 G A T A T T G A C A C T G T C C C T C T T G G C A G T G C A T T G A C T T G A A A G G T A T A T G A C T G A G C G
 T A G C A T A C A G G T T A A C C T G C A G A A A C A G T A C T T A G G T A A T T G T A G G G C G A G G A T T A A A T G A
 A A T T G C A A A A T C A C T T A G C A G C A A T G A A G A C A A T T A C C A C C A C G T G G G A G A A A T C A A A C C
 G A G C A G G G C T G T G A A A C A T G G T T G T A A T A T G C G A C T G C G A A C A C T G A A C T C T A C G C C A C T C
 C A C A A A T G A T G T T C A G G T G T C A T G G A C T G T G C C A C C A T G T A T T C A T C C A G A G T T C T T A A A
 G T T A A A G T T G C A C A T G A T G T A A G C A T G C T T C T T G A G T T T A A A T T A T G T A T A A A C A T
 A A G T T G C A T T A G A A A T C A A G C A T A A A T C A C T T C A A C T G C A
 A A A

FIGURE 8

MQRLGATLLCLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQH
KLRSAVEEEMEAEEAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGQM
VFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMLCTRSECCGDQL
CVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRLLDI
TWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFME
EVHQELEDLERSLTEEMALGEPAAAAALLGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGCGGACGCGTGGGGCTGTGAGAAAGTCCAATAAACATCATGCAACCCC
 ACGGCCACCTGTGAACTCCTCGTGCCTAGGGCTGATGTGCGTCTCCAGGGCTACTCATCC
 AAAGGCCTAATCCAACGTTCTGTCTCAATCTGCAAATCTATGGGGCCTGGGGCTCTTCTGG
 ACCCTTAACGGTACTGCCCTGGCCAATGCGCCTCGCTGGAGCCTTGCCCTCCTCTAC
 TGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCTTAATCTGCCTTCATCCGCACA
 CTCCGTTACCACACTGGTCATTGGCATTGGAGCCCTCATCCTGACCCCTGTGCAGATAGCC
 CGGGTCATCTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCCTGTAGCCGCTGC
 ATCATGTGCTGTTCAAGTGCCTCTGGTGTCTGGAAAAAATTATCAAGTTCTAAACCGC
 AATGCATACATCATGATGCCATCTACGGAAAGAATTCTGTCTCAGCCAAAATGCGTTC
 ATGCTACTCATGCGAACATTGTCAAGGGTGGCGTGGACAAAGTCACAGACCTGCTGCTG
 TTCTTGGAGCTGCTGGTGGCGAGGCGTGGGGCCTGTCCTTTCTCCGGT
 CGCATCCGGGGCTGGTAAAGACTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCATC
 ATGACCTCCATCCTGGGGCCTATGTCATGCCAGCGCTTCTCAGCGTTTGGCATGTGT
 GTGGACACGCTCTCCTGCTCCTGGAAGACCTGGAGCGAACACGGCTCCCTGGACCGG
 CCCTACTACATGTCCAAGAGCCTCTAAAGATTCTGGCAAGAAGAACGAGGCGCCCCGGAC
 AACAAAGAAGAGGAAGAAGTGACAGCTCCGGCCCTGATCCAGGACTGCACCCACCCCCACCGT
 CCAGCCATCCAACCTCACTCGCCTTACAGGTCTCCATTGTGGTAAAAAAAGGTTTAGGC
 CAGGCGCGTGGCTACGCCTGTAATCCAACACTTGAGAGGCTGAGGCGGGGGATCACCTG
 AGTCAGGAGTTGAGACCAGCCTGGCCAACATGGTGAACACTCCGTCTATTAAAAATACAA
 AAATTAGCCGAGAGTGGTGGCATGCACCTGTCACTCCAGCTACTCGGGAGGCTGAGGCAGGAG
 AATCGCTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAACC
 TGGGTGACAGACTCTGTCTCCAAAACAAACAAACAAACAAAAGATTATTAAAGATATT
 TGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLFW
TLNWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIA
RVILEYIDHKLRGVQN PVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAF
MLLMRNIVRVVVLDKVTDLLFFGKLLVVGGVGVLFFFSGRIPGLGKDFKSPHLNYYWLPI
MTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNNGSLDRPYYMSKSLLKILGKKNEAPPD
NKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCGCGCCCGGCCGGCGCCGGCGCCGAAGCCGGGAGCCACCGCC**ATG**GGGCCTGCCTGGGAG
 CCTGCTCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCTGCATCCTGTGCAGCT
 GCTGCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTCACGTTCTCCTCTTCC
 GGGTGCTGGTGTCCATCATTATGCTGAGCCCAGGGCTGGAGAGTCAGCTCTACAAGCTGCC
 GGGTGTGTGAGGAGGGGGCCGGATCCCCACCGCCTGCAGGGCACATCGACTGTGGCTCCC
 TGCTTGGCTACCGCGCTGTCTACCGCATGTGCTCGCCACGGCGCCTTCTTCTTCTT
 TCACCCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGACCCCCGGCTGCCATCCAGAATGG
 TTTGGTTCTTAAGTTCTGATCCTGGTGGCCTCACCGTGGTGCCTCTACATCCCTGAC
 GCTCCTCACCAACATCTGGTTCTACTTCGGCGTCGTGGCTCCTCCTCTTCATCCTCATCC
 AGCTGGTGCTGCTCATCGACTTGCACACTCCTGGAACCAGCGGTGGCTGGCAAGGCC
 AGTGCATTCCCCTGCTGGTACGCAGGCCTTTCTTCACTCTCCTCTACTTGCTGT
 CGATCGCGGCCGTGGCGCTGATGTTCATGTACTACACTGAGCCCAGCGGTGCCACGAG
 AGGTCTTCATCAGCCTAACCTCACCTCTGTGTCTGCGTGTCCATCGCTGCTGTCC
 AGGTCCAGGACGCCAACGCCAACTCGGGCTGCTGCAGGCCTCGGTACCGACCC
 TGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCC
 CCCAGCTGGCAACGAGACAGTTGTGGCAGGCCCGAGGGCTATGAGACCCAGTGGTGG
 CCCCAGCATTGTGGCCTCATCATCTCCTCCTGTGCACCCCTTCATCAGTCTGCGCT
 CAGACCACCGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCCCACCTATGCT
 CACAGCAGCAGCAGCAGGTTGGCAGCCTGTGAGGGCCGGCCTTGACAACGAG
 GCGTCACCTACAGCTACTCCTCTTCACTCTGCCTGGTGTGGCCTCACTGCAC
 TGACGCTCACCAACTGGTACAAGCCGGTGAGACCCGGAAGATGATCAGCAC
 GGTGGGACCGCCG
 TGTGGGTGAAGATCTGTGCCAGCTGGCAGGGCTGCTCCTCTACCTGTGG
 ACCCTGGTAGCCC
 CACTCCTCGCGAACCGCGACTTCAGC**TGA**GGCAGCCTCACAGCCTGCC
 ATCTGGTGCCTC
 CTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCC
 AGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTGCC
 CTGAGCCGGCCTTAGTCGT
 AGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAG
 CCCCCATCCCCCGCCACACCCAC
 ACGGTGGAGCTGCCTCTCCTCCCTCCCTGTTGCC
 ACTCAGCATCTCGGATGAAA
 GGGCTCCCTGTCCTCAGGCTCCAGGGAGCGGGCTGCTGGAGAGAG
 CGGGGAAC
 TCCCACC
 ACAGTGGGGCATCCGGCACTGAAGCCCTGGTGT
 CCCTGGTCACGTCCCC
 CAGGGACCC
 CTGCC
 TGTA
 AAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVES
QLYKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRDPR
AAIQNGFWFFKFLILVGLTVGAFYIPDG SFTNIWFYFGVVGSLFILQLVLLIDFAHSWNQR
WLGKAEECDRSAWYAGLFFFLLFYLLSIAAVALMFYYTEPSGCHEGVFISLNLTFCVCVS
IAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGY
ETQWWDAPSIVGLII FLLCTLFISLRSSDH RQVN SLMQTEECPPMLDATQQQQQQVAACEGRA
FDNEQDGTVTYSFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVWVKICASWAGLLL
LWTLVAPLLLNRNDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCAGGAACCACCCGTTAAGGTGTCTCTCTTAGGGATGGT
 GAGGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAG
 AACGCTCTCACCGGGAGCCAGAGCTCCCAGTGCCTCTGCGCAATATCCATTCCATCAACCCCC
 ACACAACATGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTC
 AGGAGGACTTCTGTTGTTGTCACCTTGACCTTATTGTAACATTACTGTGGATAATA
 GAGTTAAATGTGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTAC
 TATTCTCATATTTGATATATTCTCTGGCAGTTTCGATTAAAGTGTAAACTTGCA
 TATGCTGTGCAGACTGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCC
 TTTTACTAGCAAAAGTGATCCTTCGAAGCTTTCTCAAGGGCTTGGCTATGTGCTG
 CCCATCATTTCATTCATCCTGCCTGGATTGAGACGTGGTCTGGATTCAAAGTGTACCT
 CAAGAACAGAACAGAAAACAGACTCCTGATAGTTAGGATGCTTCAGAGAGGGCAGCACTT
 ATACCTGGTGGCTTGATGGTCAGTTATTCCCCTCTGAATCCGAAGCAGGATCTGAA
 GAAGCTGAAGAAAAACAGGACAGTGAGAACCACTTTAGAACTTGAGTACTACTTGTAA
 AATGTGAAAAACCCCTCACAGAAAGTCATCGAGGCAAAAGAGGCAGGCAGTGGAGTCTCCCTG
 TCGACAGTAAAGTTGAAATGGTACGTCCACTGCTGGCTTATTGAACAGCTAATAAGATT
 ATTATTGTAATACCTCACAAACGTTGATTCCATGCACATTAGTTGCCTGCCTGTGG
 CTGGTAAGGTAATGTCATGATTCCATCCTCTCTCAGTGAGACTGAGCCTGATGTGTTAACAAA
 TAGGTGAAGAAAGTCTTGTGTATTCCAATTCACAGAAATGGAATTTTTGTTCATGTCT
 CAGATTATTGTATTCTTTAACACTCTACATTCCCTGTTAACTCATGCACA
 TGTGCTTTGTACAGTTAAAAGTGTAAATAAAATCTGACATGTCAATGTGGCTAGTTTA
 TTTTCTTGTGTTGCATTATGTGTATTGCCCTGAAAGTGGACTTGCAAAAGGGAAAGAAAGG
 AATTGCGAATACATGTAAAATGTCACCAGACATTGTATTATTATCATGAAATCATGTT
 TTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTGAAATGCACAAAATGACTAAACC
 ATTCAATCATGTTCTTGCCTCAGCCAATTCAATTAAATGAACTAAATTAAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDL
LFVTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAI
ALTTAVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWIETWFLDFKVLQPQEAEENRLLIV
QDASERAALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCGACCCGCCAGGAAAGACTGA
 GGCGCGGCCCTGCCCCGCCGGCTCCCTCGGCCGCCGCCCTCCCGGGACAGAAGATGTGCT
 CCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGTGCAGGGCT
 GCCCATCCGGCTGCCAGTGCAGCCACAGACAGTCTTCTGCACTGCCGCCAGGGGACCA
 CGGTGCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTGAGAACGGCATCACCA
 TGCTCGACGCAGGCAGCTTGCCGGCCTGCCGGGCTGCAGCTCCTGGACCTGTACAGAACCC
 AGATCGCCAGCCTGCCAGCGGGTCTTCCAGCCACTGCCAACCTCAGAACCTGGACCTGA
 CGGCCAACAGGCTGCATGAAATACCAATGAGACCTCCGTGGCCTGCCGCGCTCGAGCGCC
 TCTACCTGGCAAGAACGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACC
 TCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCCGGCACTGCCCGCTGCCCTGCCCGCC
 TGCTGCTGGACCTCAGCCACAACAGCCTGCCCTGGAGCCGGCATCCTGGACACTG
 CCAACGTGGAGGCGCTGCCGGCTGGTCTGGGCTGCAGCAGCTGGACGAGGGCTTTCA
 GCCGCTTGCACACCTCCACGACCTGGATGTGTCGACAACCAGCTGGAGCGAGTGCACCTG
 TGATCCGAGGCCTCCGGGCCTGACGCCCTGCCGGCTGGAGCTGGACACACCCGATTGCC
 TGCAGGCCCTGCCTGGCGACCTCTCAGGCCCTTCCCCGCCCTGCCGCTGGCAGCTGCC
 GCAACCCCTCAACTGCGTGTGCCCTGAGCTGGTTGGCCCTGGTGCAGGAGCCACAG
 TCACACTGCCAGCCCTGAGGAGACGCCCTGCCACTTCCCACCAAGAACGCTGCCGGCTG
 TCCTGGAGCTTGACTACGCCACTTGGCTGCCAGCCACCACAGCCACAGTGCACCA
 CCACGAGGCCGTGGTGCAGGCCACAGCCTGTCTTAGCTGGCTCCTACCTGGCTTA
 GCCCCACAGGCCGGCCACTGAGGCCACAGCCTGCCACTGCCACCGACTGTAGGGC
 CTGTCCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCCTAATGGGGCACATGCCAC
 GGACACGGCACCACCTGGCGTGTGGCCCTGAAGGCTTCACGGGCTGTACTGTGAGAG
 AGATGGGGCAGGGACACGCCAGCCCTACACCAGTCACGCCGAGGCCACCACGGTCC
 CCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCCGTGGGCTGCAGCGCTACCTCC
 GGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGAACCTATCGGCCCTGATA
 GGCTGGTGACGCTGCCACTGCCCTGCCCTCGCTGAGTACACGGTACCCAGCTGCC
 ACGCCACTTACTCCGTGTGTCATGCCCTTGGGGCCGGGGGGGGGGGGGGGGGGGGGG
 CCTGCCGGGAGGCCATACACCCAGCCGTCCACTCCAACCACGCCAGTCACCCAGGCC
 GCGAGGGCAACCTGCCCTCCATTGCCGCCCTGGCCGGTGTCCCTGGCCGCGCTGG
 CTGCCGGTGGGGCAGCCTACTGTGTGCCGGGGGGGGGGGGGGGGGGGGGGGGGGGG
 AAGGGCAGGTGGGGCAGGGCTGGCCCTGGAACTGGAGGGAGTGAAGGTCCCTGGAG
 CAGGCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTGTGAGTGT
 CACTCATGGGCTCCAGGGCTGCCCTCCAGTCACCCCTCCAGCAAAGCCCTACATC
 CCAGAGAGAGACAGGGCAGCTGGGCCGGCTCTCAGCCAGTGAGATGGCCAGGCC
 CTGCCACACCACGTAAGTCTCAGTCCCAACCTCGGGATGTGTGCAGACAGGGCTGT
 CACAGCTGGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGT
 TGACGAGCCCTAACGTCCCCAGAACCGAGTGCCATTGAGGACAGTGCCCTGCC
 AACGTGCAGTCCCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCC
 TGGGCTCTCCACTCCAGGGGGACCCCTGGGGCCAGTGAAGGAAGCTCCGGAAAG
 GGAGAGCGGGTAGGCGGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAAC
 ACTGGAAAGGAAGATGCTTACGGAACATGTTGCTTTAAAATATATATTATAAGAG
 ATCCTTCCCATTATTCTGGGAAGATGTTTCAAACACTCAGAGACAAGGACTTGG
 TAAGACAAACGATGATGAAGGCCTTTGTAAGAAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLPPLLLALGPGVQGCPSPGCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFENG
 ITMLDAGSFAGLPGLQLLDSLQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRL
 ERLYLGKNRIRHIQPGAFDTLDRLLELKLDNELRALPPLRLPRLLLDDLSHNSLLALEPGIL
 DTANVEALRLAGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRI
 AQLRPEDLAGLAALQELDVSNLSQLALPGDLSGLFPRLRLLAAARNPFNCVCPLSWFGPWRE
 SHVTLASPEETRCHFPPKNAGRLLLEDYADFGCPATTTATVPTTRPVVREPTALSSSLAPT
 WLSPTAPATEAPSPPSTAPPTVGPPQPPQDCPPSTCLNGGTCHLGTRHHLACLCPEGFTGLYC
 ESQMGQGTRPSPTPVTPRPPRSLTLGIEPVSPTSLRVGLQRYLQGSSVQLRSLRLTYRNLSGP
 DKRLVTLRLPASLAEYTWTQLRPNATYSVCVMPLGPGRVPEGEEACGEAHTPPAVHSNHAPVT
 QAREGNLPLLIAPALAAVLLAALAAVGAAYCVRRGRAMAAAAQDKGQVGPAGPLELEGVKVP
 LEPGPKATEGGEALPSGSECEVPLMGFPGPGLQSPLHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
 594-600, 640-646

FIGURE 17

GCAGCGGCGAGGCGGCGGTGGCTGAGTCCGTGGCAGAGCGAAGGCGACAGCTATG
 CGGGTCCGGATAGGGCTGACGCTGCTGTGCGGTGCTGAGCTTGGCCTCGCGTCC
 TCGGATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTGACATCAGATGAG
 TCAGTAAAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTTGATTCAAGAA
 GAATCTGAATTAGAACCTCTATTCAAGAACAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAA
 AGTGTACAGAACAGATATCAGCTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAG
 CCAAAGAAAGTACGGAAACCAGCTTGACGCCATTGAAGGCACAGCACATGGGAGGCCCTGC
 CACTTCCCTTTCTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGAGGGAAAGAT
 GGCAGACTGTGGTGTGCTACAACCTATGACTACAAAGCAGATGAAAAGTGGGCTTTGTGAA
 ACTGAAGAACAGAGGCTGCTAACAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACGTGGA
 ATGAAAATCCTTAATGGAAGCAATAAGAAAAGCCAAAAAGAGAACATATCGGTATCTCCAA
 AAGGCAGCAAGCATGAACCATAACAAAGCCCTGGAGAGAGTGTATGCTCTTTATTGGT
 GATTACTTGCCACAGAATATCCAGGCAGCGAGAGAGATGTTGAGAACAGCTGACTGAGGAAGGC
 TCTCCAAGGGACAGACTGCTCTGGCTTCTGTATGCCTCTGGACTTGGTAAATTCAAGT
 CAGGCAAAGGCTTTGTATATTATAACATTGGAGCTCTGGGGCAATCTAACAGCCCACATG
 GTTTGGTAAGTAGACTTTAGTGGAAGGCTAACATAATTAAACATCAGAACAGAATTGTGGTTA
 TAGCGGCCACAACCTTTCTAGCTTCATGATCCAGATTGCTGTATTAAGACCAAATATTCA
 GTTGAACCTCCTCAAATTCTGTAAATGGATATAACACATGGAATCTACATGTAAATGAAAG
 TTGGTGGAGTCCACAATTCTTTAAATGATTAGTTGGCTGATTGCCCTAAAAAGAGAG
 ATCTGATAAAATGGCTCTTTAAATTCTCTGAGTTGGAATTGTCAGAACATCATTACAT
 TAGATTATCATAATTAAAAATTCTCTGAAATTACAAATATTGCAACAATGCCCTAACAGAACATTCA
 TGGAGTTATTGTGCAGAACATGACTCCAGAGAGCTCTACTTCTGTTTACTTTCTGATT
 GGCTGTCTCCCATTATTCTGGTCATTATTGCTAGTGACACTGTGCCTGCTCCAGTAGTC
 TCATTTCCTATTGCTAATTGTTACTTTCTTGCTAATTGGAAGATTAAACTCATTT
 TTAATAAAATTATGCTAACAGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAA

FIGURE 18

MRVRIGLTLCAVLLSLASASSDEEGSQDESLSKTTTSDESVKDHTTAGRVVAGQIFLDS
EESELESSIQEEEDSLKSQEGESVTEDISFLESPNPNKDYEETPKVRLPALTAEIGTAHGEPE
CHFPFLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAKRRQMQEAEEMMYQT
GMKILNGSNKKSQKREAYRYLQKAASMNHTKALERVSYALLFGDYLPONIQAAREMFEKLTEE
GSPKGQTALGFLYASGLGVNSSQAKALVYYTGFALGGNLIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCA GATTTA AGCCC ATTCTGCAGT GGAATT CATGAA CTAGCAAGAGGAC ACCATCTTC
 TTGTATTATA CAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAA**ATG**CTCTTTGGGTGC
 TAGGCCTCCTAATCCTCTGTGGTTCTGTGGACTCGTAAAGGAAA ACTAAAGATTGAAGACA
 TCACTGATAAGTACATTTTATCACTGGATGTGACTCGGGCTTGGAA ACTTGGCAGCCAGAA
 CTTTGATAAAAAAGGGATTCATGTAATCGCTGCCTGTACTGAATCAGGATCAACAGCCT
 TAAAGGCAGAACCTCAGAGAGACTCGTACTGTGCTCTGGATGTGACCGACCCAGAGAATG
 TCAAGAGGACTGCCAGTGGGTGAAGAACCAAGTTGGGAGAAAGGTCTCTGGGTCTGATCA
 ATAATGCTGGTGTCCCGCGTGCTGGCTCCACTGACTGGCTGACACTAGAGGACTACAGAG
 AACCTATTGAAGTGAACCTGTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTGGTCA
 AGAAAAGCTCAAGGGAGAGTTATTAAATGTCTCCAGTGTGGAGGTCGCCTGCAATCGTGGAG
 GGGGCTATACTCCATCCAAATATGCAGT GGAAGGTTCAATGACAGCTTAAGACAGGGACATGA
 AAGCTTTGGTGTGCACGTCTCATGCATTGAACCAGGATTGTTCAAAACAAACTTGGCAGATC
 CAGTAAAGGTAAATTGAAAAAAACTCGCCATTGGGAGCAGCTGTCTCCAGACATCAAACAAAC
 AATATGGAGAAGGTTACATTGAAAAAAAGTCTAGACAAACTGAAAGGCAATAATCCTATGTGA
 ACATGGACCTCTCCGGTGGTAGAGTGCATGGACCACGCTCTAACAGTCTCTCCCTAAGA
 CTCATTATGCCGCTGGAAAAGATGCCAAATTTCTGGATACCTCTGTCTCACATGCCAGCAG
 CTTGCAAGACTTTTATTGTTGAAACAGAAAGCAGAGCTGGCTAATCCCAAGGCAGTG**TGA**C
 TCAGCTAACCAAAATGTCTCCTCCAGGCTATGAAATTGGCGATTCAAGAACACATCTCCT
 TTTCAACCCATTCTTATCTGCTCCAACCTGGACTCATTTAGATCGTGTCTATTGGATTGC
 AAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGTCCCTGCTCAAGTTCTTGAAAAG
 GAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCTGTATTAGGCTTGCCTGCTGG
 TGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGATCTTACCGTGGCCTGCCCA
 TGCTTATGGTCCCCAGCATTACAGTAACCTGTGAATGTTAAGTATCATCTCTTATCTAAATA
 TAAAAAGATAAGTCAACCCAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLT
SGSTALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLINNAGVPGVLAPTDWLT
LEDYREPIEVNLFG利SVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDS
LRRDMKAFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEFYIEKSLDKLKG
NKSYVNMDLSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELAN
PKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGC GGCGGTAGC **ATG** GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTGCTCGG
 CGCACTCGCTTCCAGCACCTAACACGGACTCGGACACGGAAGGTTTCTTGGGGAAAGT
 AAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTATAAC
 AATTGACATT CAGAAATATATTCCATGCTATCAGCTTTAGCTTTATAATTCTCAGGCAG
 AGTAAATGAGCAAGCACTGAAGAAAATATTCAAATGTCAAAAAGAATGTGGTAGGTTGGTA
 CAAATTCCCGTCGTCAATT CAGATCAGATCATGACGTTAGAGAGAGGCTGCTCACAAAAACTT
 GCAGGAGCATT TTTCAAACCAAGACCTGTTCTGCTATT AACACCAAGTATAAACAGA
 AAGCTGCTCTACTCATCGACTGGAACATT CTTATATAAACCTCAAAAAGGACTTTCACAG
 GGTACCTT TAGTGGTGC CAATCTGGCATGTGAACAACTGGTTATAAAACTGTATCAGG
 TTCCTGTATGTCCACTGGTTAGCCGAGCAGTACAAACACAGCTCTAAATT TGAAGA
 AGATGGATCCTAAAGGAGGTACATAAGATAAAATGAAATGTATGCTTCATTACAAGAGGAATT
 AAAGAGTATATGCAAAAAAGTGGAGACAGTGAACAAAGCAGTAGATAAAACTAGTAAAGGATGT
 AACAGATTAAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATT CAGGCAGCAAGAGAGAA
 GAACATCCAAAAGACCCTCAGGAGAACATT TTCTTCAGGCATTACGGACCTTTCC
 AAATTCTGAATTCTTCATT CATGTGTATGTCTTTAAAAATAGACATGTTCTAAAAGTAG
 CTGTA ACTACAACCACCATCTCGATGTAGACAATCTGACCTTAATGGTAGAACACACTGA
 CATT CCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCATAAAGCCTTAGACTTAGA
 TGACAGATGGCAATTCAAGAGATCTCGGTTGTAGATAACACAAGACAAACGATCTAAAGCAA
 TACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGGCCAGAAACAGATGAAGAAAAT
 TGAAAAGATGAAGGGTTTGGTGAATATT CACGGTCTCCTACATT **TGA** TCCTTTAACCTTA
 CAAGGAGATTTTATTGGCTGATGGTAAAGCCAAACATTCTATTGTTTACTATGTT
 GAGCTACTTGCAGTAAGTCATTGTTTACTATGTTCACCTGTTGCAGTAATACACAGAT
 AACTCTTAGTGCATTACTCACAAAGTACTTTCAAACATCAGATGCTTTATTCCAAAC
 CTTTTT CACCTTCACTAAGTTGTTGAGGGGAAGGCTTACACAGACACATTCTTAGAATT
 GGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCAGTACGGGAAGACAAGTC
 AGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATTGAGACCAGTCTA
 TTAAAAAAATAAAATGGAAAAGCAAGAATAGCCTTATT TCAAATATGGAAAGAAATTATAT
 GAAAATTATCTGAGTCATTAAAATTCTCCTTAAGTGTACTTTAGAAGTACATTATGGC
 TAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCTAAAATT
 AAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLGEVKGEAKNSITDSQMDDVEVVYTIDIQK
YIPCYQLFSFYNSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHEFS
NQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMST
GFSRAVQTHSSKFFED GSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVNRLKR
EIEKRRGAQIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVS KSSCNYNH
HLDVVDNLTL MVEHTDIPEASPASTPQIIKHKA LD LDDR WQFKRSRLLDTQDKRSKANTGSSN
QDKASKMSSPETDEEIEKMKGFG EYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGAGGGCAGAGTCAGCCGAGCCAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCC
 CAAGCAGCGCGAGCGAACGCCGCCGCCACACCCCTCTGCCGTCCCCGGCGCTGCCACCCCTCCCTCC
 TTCCCCGCGTCCCCGCTCGCCGGCCAGTCAGCTGCCGGTTCGCTGCCCGCGAAACCCCGAGGTACCCAGCC
 CGCGCCTCTGCTTCCCTGGCCGCGCCCTCACGCCCTCCTCTCCCTGGCCCGCGCTGGCACCGGGG
 ACCGTTGCCTGACCGAGGGCCCAGCTACTTTGCCCGCGTCTCCCTGCCCTGCTGCCCTTCCACCAACT
 CCAACTCCTCTCCCTCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTGGCCCGCTGCCGTAGGCCCGC
 TTCCCGTCCGGTCCAAAGGTGGAACCGCGTCCGCCCGGCCGACC**ATG**GCACGGTTGGCTGCCCGCGCTT
 CTCTGCACCCCTGGCAGTGCTCAGGCCCGCTGCTGGCTGCCGAGCTAAGTCGAAAAGTTGCTCGGAAGTGC
 CGTCTTACGTGTCAAAGGCTCAACAAGAACGATGCCCTCCACGAGATCAACGGTGATCATTGAAGATC
 TGCCCCAGGGTTCTACCTGCTCTCAAGAGATGGAGGAGAACTACAGCCTGCAAAGTAAAGATGATTTCAA
 AGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTGCTTACGTTACAAGAAGTTGATGAATTCTC
 AAAGAAACTACTGAAAATGCAGAGAAATCCCTGAATGATATGTTGTGAAGACATATGGCATTATACATGCAA
 AATTCTGAGCTATTAAAGATCTCTCGTAGAGTTGAAACGTTACTACGTGGTGGAAATGTGAACCTGGAAGAA
 ATGCTAAATGACTCTGGCTCGCCTCTGGAGCGGATGTTCCGCTGGTAACTCCCAGTACCAACTTACAGAT
 GAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTCGGAGATGTCCTCGCAAATTGAAGCTC
 CAGGTTACTCGTGTGTTTAGCAGCCGTACTTCGCTCAAGGCTTAGCGGTTGGAGATGTCGTGAGCAAG
 GTCTCCGTGGTAAACCCACAGCCCAGTGTACCCATGCCCTGTTGAAGATGATCTACTGCTCCCAGTGC
 CTCGTGACTGTGAAGCCATGTTACAACATGCTCAAACATCATGAGAGGCTGTTGGCAACCAAGGGATCTC
 GATTTGAATGGAACAATTCTAGATGCTATGCTGATGGTGGCAGAGAGGGCTAGAGGGCCTTCAACATTGAA
 TCGGTGATGGATCCCATCGATGTGAAGATTCTGATGCTATTATGAACATGCAGGATAATAGTGTCAAGTGTCT
 CAGAAGGTTTCCAGGGATGTGGACCCCCCAAGCCCCTCCAGCTGGACGAATTCTCGTCCATCTGAAAGT
 GCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCAACACAGCAGCTGGCACTAGTTGGACC
 CTGGTTACTGATGTCAAGGAGAAACTGAAACAGGCCAAGAAATTCTGGCCTCCCTCGAGCAACGTTGCAAC
 GATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGAATGGAAAGGCAAAGCAGGTACCTGTT
 GCAGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAAC
 GACATACTGATCCTCGTCAAATCATGGCTCTCGAGTGATGACCAGCAAGATGAAGAATGCATA
 GACTTCTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGGAAAGTGGCTGTGAGTATCAGCAGTGC
 GACTTCTTGACTACAATGCCACTGACCATGCTGGAGAGGTGCCATGAGAAAGCCGACAGTGCTGGTGT
 GGGGCACAGGCCTACCTCCTCACTGTCTGCATCTGTTCTGGTTATGCAGAGAGAGTGGAGA**TA**ATTCTCA
 AACTCTGAGAAAAAGTGTCAATCAAAAGTAAAGGCACCAAGTATCAGTTCTACCATCCTAGTGACTTGC
 TTTTAAATGAATGGACAACAATGTACAGTTTACTATGTGGCACTGGTTAAGAAGTGCTGACTTGT
 TCATTCACTGGGAGGAAAGGGACTGTGCATTGAGTTGGCTCTGCTCCCCAAACCATGTTAACGTGGCT
 AACAGTGAGGTACAGAACTATAGTTAGTGCGATTGTGATCTCGCCTGTTACAAGCAAACCAGGGTCC
 TTCTCATTGCTTGTGGTTTTCTCAACTGTGATCTCGCCTGTTACAAGCAAACCAGGGTCC
 CTTGGCACGTAACATGTACGTATTCTGAAATATTAAATAGCTGTACAGAACAGGTTATTATCATGTT
 TTATTAAAAGAAAAAGCCAAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRLYVSKGFNKNDAPLHEINGDHLKICPQG
STCCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKS LNDMFVK
TYGHLYMQNSELFKDLFVELKRYVVGVVNLEEMLNDFWARLLERMFR LVNSQYHFTDEYLEC
VSKYTEQLKPFGDVPRKLKLQVTRAFVAARTFAQGLAVAGDV VSKVSVNPTAQCTHALLKMI
YCSHC RGLVTVKPCNYCSNIMRGCLANQGDLDFEWNNFIDAML MVAERLEGPFNIESVMDPI
DVKISDAIMNMQDNSVQVSQKV FQCGPPKPLPAGRISRSI SEAFSARFRPHPEERPTAA
GTSLDRLVTDVKEKLKQAKKF WSSLPSNCNDERMAAGNGNE DDCWNGKGKS RYLFAVTGNGL
ANQGNNPEVQV DTSKP DILILRQIMALRVMTSKMKNAYNGNDV DFFDISDESSGE GSGCEY
QQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTGCCCTCAAATGGAACGCTGGCCTGGACTAAAGCATAGACCACCAAGGCTGAGTATCCTG
ACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCATTATATTCTTCAAGC
AACTTACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTCCCTCCTGTTGCTGC
CACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGTCGCCAGAGGCCACAGGG
ACCGAGGCCAGGCTTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATT
GGTCCTGAGAGCCCCGAGAAGAAAATTGACAGTGTCTGGGCTGCCAAAGAACAGCAGTGCC
CCTGTGATCATTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAA
ACAAGCATTCCAGAGCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAACAGCTTGTCTGC
CTTGTAGGAGCTCTGAGCGCCCACCTCCAATTAAACATTCTAGCCAAGAACAGTGAG
CACACCTACCAGACACTCTTCTTCCCACCTCACTCTCCACTGTACCCACCCCTAAATCAT
TCCAGTGCTCTCAAAAAGCATGTTTCAAGATCATTGTTGCTCTAGTGTCT
CTTCTCTCGTCAGTCTAGCCTGTGCCCTCCCTACCCAGGCTTAGGCTTAATTACCTGAAA
GATTCCAGGAAACTGTAGCTCCTAGCTAGTGTCAATTAAACCTTAAATGCAATCAGGAAAGTA
GCAAACAGAAGTCAATAATATTAAATGTCAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRK
FMTVSGLPKKQCPDCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCAAGCGGAAGCACAGCTC
 AGAGCTGGTCTGCCATGGACATCCTGGTCCCACTCCTGCAGCTGCTGGTGTGCTTACCC
 TGCCCCTGCACCTCATGGCTCTGCTGGCTGGCAGCCCCGTGCAAAAGCTACTTCCCCT
 ACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAACGGGAGCTCT
 TCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGCTGCG
 GAACCGGAGCCAACTTCAGTTCTACCCACCAGGCTGCAGGGCACCTGCCTAGACCCAAATC
 CCCACTTGAGAAGTTCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGT
 TTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCT
 GCACTCTGGTGTGCTGTGCAGAGCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTAC
 TGAGACCGGGAGGTGTGCTCTTCTGGAGACATGTGGCTGAGAACACCATAATGGAAGCTGGCCT
 TCATGTGGCAGCAAGTTCTGAGCCCACCTGGAAACACATTGGGATGGCTGCCTCACCA
 GAGAGACCTGGAAGGATCTTGAGAACGCCAGTTCTCCGAAATCAAATGGAACGACAGCCCC
 CTCCCTTGAAGTGGTACCTGTTGGCCCCACATCATGGAAAGGCTGTCAAACAATCTTCC
 CAAGCTCCAAGGCACTCATTGCTCCTCCCCAGCCTCCAATTAGAACAGCCACCCACCAGC
 CTATCTATCTTCACTGAGAGGGACCTAGCAGAATGAGAGAACATTGATGTACCACTACT
 AGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCAACTCAATCCCGCCTCGACAGTGA
 AAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCCCTGTTGTATCCTCAACTG
 CAAGTTCTGGACTAGTCTCCAACGTTGCCTCCCAATGTTGTCCCTTCCTCGTTCCCAT
 GGTAAAGCTCCTCTCGCTTCCCTGAGGCTACACCCATGCGTCTCTAGGAACGGTCACAA
 AAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCCTCTCCCCACTACCACCTCTT
 CCTGAGCTGGGGCACCAAGGGAGAATCAGAGATGCTGGGATGCCAGAGCAAGACTCAAAGAG
 GCAGAGGTTTGTCTCAAATATTTTAATAAATAGACGAAACCACG

FIGURE 28

MDILVPLLQLLVLLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIK
GLTGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAP
GEDMRQLADGSMDVVVCTLVLCVQSPRKVLQEVRVRLPGGVLFWEHVAEPYGSWAFMWQQ
VFEPTWKHIGDGCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFPSKA
LICSFPSLQLEQATHQPIYPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTGCCTATCCACCTCCCCAAGCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTG
CTGCTGCTGCTGCTAAAGGCTCATGCTTGGAGTGGGACTGGTCGGGCCAGAAAGTCTCT
TCTGCCACTGACGCCCATCAGGGATTGGGCCTTCTTCCCCCTCCTTCTGTGTCTCCTG
CCTCATCGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGGAGAAAGTGGG
GATGGCTAAGAAAGCTGGAGATAGGAAACAGAAGAGGGTAGTGGGTGGCTAGGGGGCTGC
CTTATTAAAGTGGTTTTATGATTCTTATACTAATTATAAAAGATATTAAGGCCCTGTT
CATTAAGAAATTGTTCCCTTCCCCTGTGTTCAATGTTGAAAGATTGTTCTGTGAAATATG
TCTTATAATAAACAGTTAAAAGCTGAAAAA

FIGURE 30

MLLLTLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTC SQAQP
RGEGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCTTCAACTATAACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAG
 TTCCTCCAAGCAAGTCATTTCCCTATTAAACCGATGTGTCCCTCAAACACACCTGAGTGCTACT
 CCCTATTGCATCTGTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAAATC
TGCGGGAAAGAGATAACAATCCTGGCCTGTGTATCCTCGCATTAGCCTGTCTTGGCCATGA
 TGTTTACCTTCAGATTCATCACCAACCCTCTGGTTCACATTTCATTGGTTATTTGG
 GATTGTTGTTGTCTCGGGTGTATGGTGGCTGTATTATGACTATAACCAACGACCTCAGCA
 TAGAATTGGACACAGAAAGGGAAAATATGAAGTGCCTGCTGGGTTGCTATCGTATCCACAG
 GCATCACGGCAGTGCTCGTCTGATTTCAGAAAGAGAATAAAATTGACAGTTG
 AGCTTTCAAATCACAAATAAACGCCATCAGCAGTGCTCCCTGCTGTTCCAGCCACTGT
 GGACATTGCCATCCTCATTTCCTGGGTCCTCTGGGTGGCTGTGCTGAGCCTGGAA
 CTGCAGGAGCTGCCAGGTTATGGAAGGCGGCCAAGTGGAAATATAAGCCCCTTCGGCATT
 GGTACATGTGGCGTACCATTAATTGGCCTCATCTGGACTAGTGAATTCATCCTGCGTGCC
 AGCAAATGACTATAGCTGGGCAGTGGTTACTTGTATTTAACAGAAGTAAAAATGATCCTC
 CTGATCATCCCATCCTTCGCTCTCTCCATTCTCTTCTACCATCAAGGAACCGTTGTGA
 AAGGGTCATTTAATCTCTGTGGTAGGATTCCGAGAATCATTGTCATGTACATGCAAAACG
 CACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGATGCTACTGCTGTT
 TCTGGTGTCTGACAAATACCTGCTCCATCTCAACCAGAATGCATATACTACAACGTCTATT
 ATGGGACAGATTCTGTACATCAGAAAAGATGCATTCAAAGAAACTCAAGTC
 ACTTTACATCTATTAACTGCTTGAGACTTCATAATTTCAGGAAAGGTGTTAGTGGTGT
 GTTCACTGTTGGAGGACTCATGGCTTTAACTACAATCGGGCATTCCAGGTGTGGCAG
 TCCCTCTGTTATTGGTAGCTTTGCCTACTTAGTACGCCATAGTTTATCTGTGTTG
 AAACTGTGCTGGATGCACCTTCCTGTGTTGCTGTTGATCTGGAAACAAATGATGGATCGT
 CAGAAAAGCCCTACTTATGGATCAAGAATTCTGAGTTCTGAAAAAGGAGCAACAAATTAA
 ACAATGCAAGGGCACAGCAGGACAAGCACTCATTAGGAATGAGGAGGGAACAGAACTCCAGG
 CCATTGTGAGATAGATACCCATTAGGTATCTGTACCTGGAAAACATTCTAAGAGCCA
 TTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTAAAGACCTAA
 TAAACCCTATTCTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITTLLVHIFISLVLGLLFVCGLWWLYDYTNDLS
IELDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISAPFLLFQPL
WTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILAC
QQMTIAGAVVTCYFNRSKNPDPDHPILSSLSILFFYHQGTVVKGFLISVVRIPRIIVMYMQN
ALKEQQHGALSRYLFRCYCFCFWCLDKYLLHLNQNAYTTAINGTDFCTSAKDAFKILSKNSS
HFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLVAFFAYLVAHSFLSVF
ETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQ
AIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTGGACCTCTCCCTGTTCTCCTTA
 GAATAATTGTATGGGATTGTGATGCAGGAAAGCCTAACGGAAAAAGAATATTCAATTCTGTG
 TGGTAAAAATTGGGAAAGGGAAAGGGAAAGGGTGTCAATTGTGATATT**AT**
GAGGACTGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCTGTTGCTGGTGAC
 TGGAGTACATTCAAACAAAGAAACGGCAAAGAAGATTAAAAGGCCAAGTTCACTGTGCCTCA
 GATCAACTGCGATGTCAAAGCCGAAAGATCATCGATCCTGAGTTATTGTGAAATGTCCAGC
 AGGATGCCAAGACCCAAATACCATGTTATGGCACTGACGTGTATGCATCCTACTCCAGTGT
 GTGTGGCGCTGCCGTACACAGTGGTGTGCTGATAATTCAAGGAGGGAAAATACTGTTGGAA
 GGTTGCTGGACAGTCTGGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACC
 ACGATGGAGAGAATCCTTATCGTCTAGAAAGTAAACCCAAAAGGGTGTAACTACCCATC
 AGCTCTTACATACATCATCGAAAAGTCCAGCTGCCAAGCAGGTGAGACCACAAAAGCCTA
 TCAGAGGCCACCTATTCCAGGGACAACGCACAGGCCGGTCACTGTGATGCAGCTCTGGCTGT
 CACTGTAGCTGGCCACCCCCCACCTGCCAAGGCCATCCCCTCTGCTGCTTCTACCAC
 CAGCATCCCCAGACCACAATCAGTGGGCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGC
 CACCTACACAAGCAGCCAAAACAGGCCAGAGCTGATCCAGGTATCAAAGGCAAGATCCTTC
 AGGAGCTGCCTCCAGAAACCTGTTGGAGCGGATGTCAGCCTGGACTTGTCCAAAAGAAGA
 ATTGAGCACACAGTCTTGGAGCCAGTATCCCTGGGAGATCCAAACTGCAAATTGACTTGTC
 GTTTTAATTGATGGGAGCACCAGCATTGGCAAACGGGATTCCGAATCCAGAAGCAGCTCCT
 GGCTGATGTTGCCAAGCTCTGACATTGGCCCTGCCGGTCCACTGATGGGTGTTGTCAGTA
 TGGAGACAACCCTGCTACTCACTTAACCTCAAGACACACAGAATTCTCGAGATCTGAAGAC
 AGCCATAGAGAAAATTACTCAGAGAGGGACTTCTAATGTAGGTGGGCCATCTCCTTGT
 GACCAAGAACTCTTCCAAAGCCAATGGAAACAGAAGCAGGGCTCCAAATGTGGTGGTGGT
 GATGGTGGATGGCTGGCCACGGACAAAGTGGAGGGAGGCTCAAGACTTGCAGAGAGTCAGG
 AATCAACATTTCTTCATCACCATTGAAGGTGCTGCTGAAATGAGAAGCAGTATGTGGTGG
 GCCCAACTTGCACAAAGGCCGTGCAAGAACAAACGGCTTCACTCGCTCCACGTGCAGAG
 CTGGTTGGCCTCCACAAGACCCCTGCAGCCTCTGGTGAAGCAGGGCTGCGACACTGACCGCCT
 GGCCTGCAGCAAGACCTGCTGAACCTGGCTGACATTGGCTCGTCATCGACGGCTCCAGCAG
 TGTGGGACGGCAACTCCGCACCGTCCTCCAGTTGTGACCAACCTCACCAAAAGAGTTGA
 GATTTCGACACGGACACGCGCATCGGGCGTGCAGTACACCTACGAACAGCAGGGCTGGAGTT
 TGGGTTGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGCTACTG
 GAGTGGTGGCACCAGCACGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTCTCAAGAAGTC
 CAAGCCCAACAAGAGGAAGTTAATGATCCTCATCACCGACAGGGAGGTCTACGACGACGTCCG
 GATCCCAGCCATGGCTGCCATCTGAAGGGAGTGATCACCTATGCATAGGCAGTGGCTGG
 TGCCCAAGAGGGACTAGAAGTCATTGCCACTCACCCGCCAGAGACCACCTCCTTGTGGA
 CGAGTTGACAACCTCCATCAGTATGTCCCCAGGATCATCCAGAACATTGTACAGAGTTCAA
 CTCACAGCCTCGGAACTGAATTCAAGAGCAGGCAGAGCACCAGCAAGTGCTGCTTFACTA
 ACGTGGTGGACCACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTGGCAGGGCA
 TGGAGAAACAAATGTCTGTTATTATTCTTGCCTCATGCTTTCATATTCCAAAACCTGG
 AGTTACAAAGATGATCACAAACGTATAGAATGAGCAGGAAAGGCTACATCATGTTGAGGGG
 GGAGATTTACATTGACAATTGTTCAAAATAATGTTGGAAATACAGTGCAGCCCTTAC
 GACAGGCTTACGTAGAGCTTGTGAGATTAAAGTTGTTATTCTGATTGAACTCTGTAA
 CCCTCAGCAAGTTCATTTGTGACATGACAATGTAGGAATTGCTGAATTAAATGTTAGAAGG
 ATGAAAAATAAAG
 AAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCP
AGCQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSL
PRWRESFIVLESKPKKGVTYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLA
VTVAVATPTTLPKPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQNRPRADPGIQRQDP
SGAAFKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQL
LADVAQALDIGPAGPLMGVVQYGDNPATHFNLKTHTNSRDLKTAIEKITQRGGLSNVGRAISF
VTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVV
EPNFANKAVCRTNGFYSLHVQSWFGLHKTLQPLVKRVCDTDRLACSKTCLNSADIGFVIDGSS
SVGTGNFRTVLQFVTNLTKFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGY
WSGGTSTGAAINFQLEQLFKKSVPNKRKLMILITDGRSYDDVRIPAMAHLKGVITYAIGVAW
AAQEELEVIATHPARDHSFFVDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTAGGAGGTGGCTGCGTTGGAAAAGCTATCAAGGAA
 GAAATTGCCAACCATGTCTTTCTGTTTCAGAGTAGTCACAACAGATCTGAGTGT
 TAATTAAAGCATGGAATACAGAAAACAACAAAAACTTAAGCTTAATTCTGGAATTCCA
 CAGTTTCTTAGCTCCCTGGACCCGGTTGACCTGTTGGCTCTCCGCTGGCTGCTCTAC
 GTGGTGCCTCCGACTACTCACCCCGAGTGTAAAGAACCTCAGCTCGCGTCTGAGCTG
CTGTGGATGGCCTCGGCTCTGGACTGTCCTCCGAGTAGGATGTCAGATCCCTCAA
 TGGAGCCTCCTGCTGTCACTCCTGAGTTCTTGATGTGGTACCTCAGCCTCCCCAC
 TACAATGTGATAGAACCGTGAACGGATGTACTTCTATGAGTATGAGCCGATTACAGACAA
 GACTTCACTCACACTCGAGAGCATTCAAACGTCTCATCAAAATCCATTCTGGTCATT
 CTGGTACCTCCCACCCTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGTGAA
 AAAAGTCTTGGTGGGATATGAGGTTCTACATTCTATTAGGCCAAGAGGGCTGAAAAG
 GAAGACAAAATGTTGGCATTGTCCTAGAGGATGAAACACCTTATGGTACATAATCCGA
 CAAGATTTTAGACACATATAAACCTGACCTGAAAACCATTATGGCATTAGGTGGTA
 ACTGAGTTTGCCTAACATGCCAAGTACGTAATGAAGACAGACACTGATGTTTCAATACT
 GGCAATTAGTGAAGTATCTTAAACCTAACCAACTCAGAGAAGTTTACAGGTTATCCT
 CTAATTGATAATTATTCTATAGAGGATTACCAAAAAACCCATATTCTACCAGGAGTAT
 CCTTCAAGGTGTTCCCTCCACTGCAGTGGGTTGGGTATATAATGTCCAGAGATTGGT
 CCAAGGATCTATGAAATGATGGTCACGTAAAACCCATCAAGTTGAAGATGTTATGTC
 ATCTGTTGAATTATTAAAAGTGAACATTCAATTCCAGAAGACACAAATCTTCTTCTA
 TATAGAATCCATTGGATGTCTGTCAACTGAGACGTGATTGCAGCCCATTGGCTTCT
 AAGGAGATCATCACTTTGGCAGGTATGCTAACAGGAAACACCACATGCCATTTTAACTTCAC
 ATTCTACAAAAAGCCTAGAAGGACAGGATACCTGTGGAAAGTGTAAATAAGTAGGTACTG
 TGGAAAATTGTCATGGGAGGTACACTGTGCTGGCTTACACTGAACATGAAACTCATG
 GACTGGAGACTGGAGGGTTACACTGTGATTAGTCAGGCCCTCAAAGATGATATGTGG
 AGGAATTAAATATAAGGAATTGGAGGTTTGCTAAAGAAATTAAATAGGACAAACAATTG
 GACATGTCATTCTGTAGACTAGAATTCTTAAAGGGTGTACTGAGTTATAAGCTCACT
 CTGTAACAAACAAACATGTAGAGTTTATTGAACATGTAGTCACCTGAAGGTTTGT
 GTATATCTTATGTGGATTACCAATTAAAAATATGTAGTTCTGTGTCAAAACTTCT
 CTGAAGTTATACTGAACAAATTACCTGTTGGCTTACATGCAAACATTCCAGTTAC
 GCAGTATTACAGTTATTATTAAAATTACTCAACTTGTGTTTAAATGTTGAC
 GATTCAATACAAGATAAAAGGATAGTGAATCATTCTTACATGCAAACATTCCAGTTAC
 TTAACGTGATCAGTTATTATTGATAACATCACTCCATTAAATGTAAAGTCAGGTCA
 ATATCAGTAATCTCTGGACTTGTAAATATTACTGTGGTAATATAGAGAAGAATTAAAG
 CAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDF
HFTLREHSNC SHQN PFLVILVTSH PSDV KARQ AIRT WGEK KS WGYEV LTF LLG QEA EKED
KMLALSLEDEHLLYGDII RQDFLDTYNNLT LKTIMA FRW VTEFCPNA KYVMKT DVF INTGN
LVKYLLNL NHSEK FFTGYPLIDNYSYRGFYQK THISYQEYPFKVFP PYCS GLGYIMSRDLVPR
IYEMMGHV KPIKFEDV YVGICLNLLKV NIHI PEDTNL FFLYRI HLDVCQLRRVIAAHGFSSKE
IITFWQVMLRNTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCGGCAAGGATGGGAGCTGGTTGCTGGACGCAGTGGGGCTCACCTTCAGCTCCT
TCTCATCTCGCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTG
TCGGGAGTGCTGTGAATATGATCAGATTGAGTGCCTGCCCTGGAAAGAGGGAAAGTCGTGGTTATACCATCCC
TTGCTGCAGGAATGAGGAGAATGAGTGTGACTCCTGCCCTGATCCACCCAGGTTGATCCATCTTGAAAAGTCAA
GAGCTGCCGAAATGGCTCATGGGGGGTACCTTGGATGACTCTATGTGAAGGGTTCTACTGTGCAGAGTGCCG
AGCAGGCTGGTACGGAGGAGACTGCATGCGATGTGCCAGGTTCTCGAGCCCCAAAGGGTCAGATTTGTTGGA
AAGCTATCCCTAAATGCTCACTGTGAATGGACCATTGCTAAACCTGGGTTGTCATCCAACTAAGATTG
CATGTTGAGTCTGGAGTTGACTACATGTGCCAGTATGACTATGTTGAGGTTGATGGAGACAACCGCGATGG
CCAGATCATCAAGCGTGTCTGGCAACGAGCGGCCAGCTCTATCCAGAGCATAGGATCCTCACTCCACGTCT
CTTCCACTCCGATGGCTCCAAGAATTGACGGTTCCATGCCATTATGAGGAGATCACAGCATGCTCCTCATC
CCCTTGTTCATGACGGCACGTGCCCTTGACAAGGCTGGATCTTACAAGTGTGCCCTGCTGGCAGGCTATAC
TGGGCAGCGCTGTGAAAATCTCCTTGAAGAAAGAAACTGCTCAGACCCCTGGGGCCAGTCATGGTACCAAGAA
AATAACAGGGGCCCTGGCTTATCAACGGACGCCATGCTAAAATTGGCACCGTGGTGTCTTCTTTGTAACAA
CTCCTATGTTCTTAGTGGCAATGAGAAAAGAACCTGCCAGCAGAACATGGAGAGTGGCAGGGAAACAGCCATCTG
CATAAAAGCCTGCCGAGAACCAAAGATTTCAGACCTGGTGAGAACGGAGAGTTCTCCGATGCAGGTTAGTCAG
GGAGACACCATTACACCAGCTATACTCAGCGCCCTCAGCAAGCAGAAACTGCAGAGTGCCTACCAAGAACCC
AGCCCTCCCTTGGAGATCTGCCATGGGATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACC
CTTCTACCGCCGCCTGGCAGCAGCAGGAGACATGTCAGGACTGGGAAGTGGAGTGGCAGGCAACATCCTG
CATCCCTATCTGCCGAAAATTGAGAACATCACTGCTCCAAAGACCCAGGGTTGCCCTGGCAGGCAAGAC
CATCTACAGGAGGACCAGCGGGTGCATGACGGCAGCCTACACAAGGGAGCGTGGTCTAGTCTGCAGCGGTGC
CCTGGTGAATGAGCGACTGTGGTGGCTGCCACTGTGTTACTGACCTGGGAAGGTACCATGATCAAGAC
AGCAGACCTGAAAGTTGGGAAATTCTACCGGGATGATGACCGGGATGAGAACCCATCCAGAGCCTACA
GATTCTGCTATCATTGCTACCCAACTATGACCCATCCTGCTGATGCTGACATGCCATCCTGAAGCTCCT
AGACAAGGCCGTATCAGCACCCGAGTCCAGCCATCTGCCCTGCCAGTCGGATCTCAGCACTCCCTCCA
GGAGTCCCACATCACTGTGGCTGGATGTCCTGGCAGACGTGAGGAGCCCTGGCTCAAGAACGACACACT
GCGCTCTGGGGTGGTCAGTGTGGACTCGCTGCTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAG
TGTCACTGATAACATGTTCTGTGCCAGCTGGAACCCACTGCCCTCTGATATCTGCACTGCAGAGACAGGAGG
CATCGCGCTGTGTCCTCCGGACGAGCATCCTGAGCCACGCTGGCATCTGATGGACTGGCAGCTGGAG
CTATGATAAAACATGCCACAGGCTCTCACTGCCCTACCAAGGTGCTGCCCTTAAAGACTGGATTGAAAG
AAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAGTGTATATCCGTCTGTACGTGTGTCATTGCG
TGAAGCAGTGTGGCCTGAAGTGTGATTGGCCTGTGAACCTGGCTGTGCCAGGGCTCTGACTCAGGGACAAA
ACTCAGTGAAGGGTAGACCTCCATTGCTGGTAGGCTGATGCCCGTCCACTACTAGGACAGCCAATTGGAA
GATGCCAGGGCTTGCAAGAAGTAAGTTCTCAAAGAACCATATAACAAACCTCTCCACTCCACTGACCTGGT
GGTCTCCCCAACTTCAGTTACCAATGCCATCAGCTGACCGAGGGAAAGATCTGGCTCATGAGGCCCTT
TGAGGCTCTCAAGTTAGAGAGCTGCCCTGTGGGACAGCCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCC
GTGTACATGCCACAGTACAGTCTGGCCTTCCCTCCATCTCTGTACACATTAAATAAAAGGGTTG
GCTTCTGAACACAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAA
AAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAA

FIGURE 38

MELGCWTQLGLTFLQLLISSLPREYTVINEACPGAEWNIMCRECCEYDQIECVCPGKREVVG
 YTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCM
 RCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDN
 RDGQIIKRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVL
 DKAGSYKCACLAGYTQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFC
 NNSYVLSGNEKRTCQQNGEWSGKQPICIKACREP KISDLVRRVLPMQQSRETPLHQLYSAA
 FSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRTCLRTGKWSGRAPS
 CIPICGKIENITAPKTQGLRW PWQAAIYRRTSGVHD GSLHKGA WFLVC SGALVNERTVVVAH
 CVTDLGKV TMK TADLK VVLGKF YR DDRDE KT IQSLQ I SAI ILHP NYD PILL DADIA I LKLL
 DKARI STRV QPI CLA ASR DLST SFQ ESH ITV AGW NVI AD VR SP GF KND TL RG VV SV DS LLC
 EEQHEDHGIPVSVTDNMFCASWEPTAPS DICTAETGGIAAVSF PGRASPEPRW HLMGLVWSY
 DKTCSHRLSTAFTKVLPKDWIERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
 474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTCCTACATCCTCATCTGAGAATCAGAGAGCATAATCTTCTACGGGCCGTGATTATTAACGTGGCTTA
 ATCTGAAGGTTCTCAGTCAAATTCTTGTATCTACTGATTGTGGGGCATGGCAAGGTTGCTAAAGGAGCTT
 GGCTGGTTGGCCCTTGTAGCTGACAGAAGGTGCCAGGGAGAATGCAGCACACTGCTGGAGAATGAAGGCGC
 TTCTGTTGCTGGCTTGCTTAGCTGCTAACACTACATTGACAATGTGGCAACCTGCACCTCCTGTATT
 CAGAACTCTGTAAAGGTGCCTCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTACAAGATGGCTGTCCAG
 ACGGCTGTGCAGCCTCACAGCCACGGCTCCCTCCCCAGAGGTTCTGCAGCTGCCACCATCTCCTTAATGACAG
 ACGAGCCTGGCCTAGACAACCCCTGCCTACGTGTCCCTGGCAGAGGACGGCAGCCAGCAATCAGCCCAGTGGACT
 CTGGCCGGAGCAACCGAACTAGGGCACGGCCCTTGAGAGAGATCCACTATTAGAACAGATCATTAAAAAAATAA
 ATCGAGCTTGAGTGTCTCGAAGGACAAAGAGCGGGAGTGCAGTTGCCAACATGCCACCAGGGCAGGGAAA
 ATTCTGAAAACACCACTGCCCTGAAGTCTTCCAAGGTTGTACCACTGATTCCAGATGGTAAATTACCAAGCA
 TCAAGATCAATCGAGTAGATCCCAGTGAAAGCCTCTATTAGGCTGGTGGAGGTAGCGAAACCCACTGGTCC
 ATATCATTATCCAACACATTTATCGTATGGGTGATGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTC
 TAAAGGTCAACGGATGGACATCAGCAATGTCCCTACAACACTACGCTGTGCGTCTCCTGCCAGCCCTGCCAGG
 TGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCCGCAGCAGGAACAATGGACAGGGCCCGGATGCCCTACAGAC
 CCCGAGATGACAGCTTCATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGAATAAAACTGGTGCAGCA
 AGGTGGATGAGCCTGGGTTTCATCTCAATGTGCTGGATGGCGGTGTGGCATATCGACATGGTCAAGCTTGAGG
 AGAATGACCGTGTGTTAGCCATCAATGGACATGATCTCGATATGGCAGCCCAGAAAGTGCAGCTCATCTGATT
 AGGCCAGTGAAAGACGTGTTACCTCGTGTCCGCCAGGTTGGCAGCGGAGCCCTGACATCTTCAGGAAG
 CCAGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCAGGGAGAGGAACAACACTCCAAAGCCCCTCCATCCTA
 CAATTACTTGTATGAGAAGGTGGTAAATATCCAAAAGACCCGGTGAATCTCGGATGACCGTGCAGGG
 GAGCATCACATAGAGAATGGGATTGCCTATCTATGTATCAGTGTGAGGCCGGAGGAGTCATAAGCAGAGATG
 GAAGAATAAAACAGGTGACATTTGTTGAATGTGGATGGGTCGAACGTGACAGAGGTGAGCCGGAGTGAGGCAG
 TGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTGGAAAGTCAAAGAGTATGAGCCCAGGAAG
 ACTGCAGCAGCCCAGCAGCCCTGGACTCCAACCACAACATGGCCCACCCAGTGACTGGTCCCCATCCTGGTCA
 TGTGGCTGGAATTACCACGGTGTGTATAACTGTAAAGATATTGTATTACGAAGAACACAGCTGGAAGTCTGG
 GCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAAACAAACCTTTTCTCAAATCCATTGTTGAAGGAA
 CACCAGCATACAATGATGGAAGAATTAGATGTGGTGTATTCTTCTGCTGCAATGGTAGAAGTACATCAGGAA
 TGATACATGCTGCTGGCAAGACTGCTGAAAGAACTAAAGGAAGAATTACTCTAACTATTGTTCTGGCCTG
 GCACCTTTTATAGAATCAATGATGGTCAGAGGAAACAGAAAAATCAAATAGGCTAAGAAGTGTGAAACACT
 ATATTATCTTGTCAAGTTTATTTAAAGAAAGAACATATTGTAAGGAAACTACTAGTTTTTCAGTGTGGAGGAT
 TTCTCATTACTCTACAACATTGTTATTTCTATTCAATAAAAGCCCTAAAACAACAAATGATTGATT
 TGTATACCCCACTGAATTCAAGCTGATTAAATTAAAGGTTATGCTGAAGTCTGCCAAGGGTACATTAT
 GGCCATTTTAATTACAGCTAAATATTAAAGGTTATGCTGAGAACGTTGCTTTCATCAAACAAGAAT
 AAATATTTCAGAAGTTAAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTATA
PSPEVSAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDGRSNRTRARPFERSTIRSRSFK
KINRALSVLRRTKSGSAVANHADQGRESENTTAPEVFPRLYHLIPDGEITSIKINRVDPSES
LSIRLVGGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPC
QVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVIFNV
LDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRSPDIFQE
AGWNSNGSWSPGPERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYV
ISVEPGGVISRDGRIKTGDILLNDGVELTEVRSEAVALLKRTSSIVLKALEVKEYEPQED
CSSPAALDSNHNMAPPSDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGN
KPFFIKSIVEGTPAYNDGRIRCGRDILLAVNGRSTSGMIHACLRLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAA
 GCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTACCGTAATAAAAAACATGGGCTT
 CAACCTGACTTCCACCTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTGTGCCTGAC
 AGTGGTTGGGTGGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCTAAAGCAA
 GGAGTTCATGGCTAATTCCATAAGACCCTCATTGGGGAAAGGGAAAAACTCTGACTAATGA
 AGCATCCACGAAGAAGGTAGAACTTGACAACACTGTCCTCTGTCTCCTTACCTCAGAGGCCA
 GAGCAAGCTCATTTCAAACCAGATCTCACTTGGAAAGAGGTACAGGCAGAAAATCCCAG
 GTCCAGAGGCCGGTATGCCCTCAGGAATGTAAAGCTTACAGAGGGTGCCTACCTCGTTCC
 CCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTGCAGAG
 GCAGCAGCTGGATTATGGCATCTACGTCACTCCACCAGGCTGAAGGTAAAAAGTTAAC
 CAAACTCTGAATGTGGCTATCTAGAACCCCTCAAGGAAGAAAATTGGACTGCTTATATT
 CCACGATGTGGACCTGGTACCCGAGAATGACTTAACTTACAAGTGTGAGGAGCATCCAA
 GCATCTGGTGGTGGCAGGAACAGCACTGGTACAGGTTACAGTGGATATTGGGG
 TGTTACTGCCCTAACGAGAGAGCAGTTCAAGGTGAATGGATTCTAACAAACTACTGGGG
 ATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTCCCG
 GCCCCTGCCTGAAGTGGTAAATATAACATGGCTTCCACACTAGAGACAAAGGAATGAGG
 GAACGCAGAACGGATGAAGCTCTACACCAAGTGTACGAGTCTGGAGAACAGATGGTTGAG
 TAGTTGTTCTATAAATTAGTATCTGTGGAACACAATCCTTATATACACATCACAGTGG
 TTTCTGGTTGGTGCATGACCCTGGATCTTGGTATGTTGGAAGAACTGATTCTTGT
 GCAATAATTGGCTAGAGACTTCAAATAGTAGCACACATTAAGAACCTGTTACAGCTCATT
 GTTGAGCTGAATTTCCTTTGTATTTCTTAGCAGAGCTCCTGGTATGTAGAGTATAAA
 ACAGTTGTAACAAGACAGCTTCTAGTCATTGATCATGAGGGTAAATATTGTAATATGG
 ATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAAAATGAACGCTATTGAGGACT
 CTGGTTGAAGGAGATTATTAAATTTGAAGTAATATATTATGGATAAAAGGCCACAGGAAA
 TAAGACTGCTGAATGTCTGAGAGAACCAAGAGTTGTTCTCGTCCAAGGTAGAAAGGTACGAAGA
 TACAATACTGTTATTCAATTCTGTACAATCATCTGTGAAGTGGTGGTGTAGGTGAGAAG
 GCGTCCACAAAGAGGGAGAAAAGGCACGAATCAGGACACAGTGAACCTGGGAATGAAGAG
 GTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTGCAGGTGCTGA
 TAGCCTTCAGGGAGGACCTGCCAGGTATGCCTCCAGTGATGCCACCAGAGAACATTC
 TCTATTAGTTAAAGAGTTGTAAAATGATTTGTACAAGTAGGATATGAATTAGCAGT
 TTACAAGTTACATATTAACATAATAAAATGTCTATCAAATACCTCTGTAGTAAAATGTG
 AAAAGCAAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLGKGKTL
TNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGYRPQECKALQRVAI
LVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDC
FIFHDVDLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGTALSREQFFKVNGFSNN
YWGWGGEDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTD
GLSSCSYKLVSVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

FIGURE 44

MALSSQIWAACLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRDTHF
PICIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCACTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCAC
CCTCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGT
CGGTTCCGTTGGTGGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAAGCAAGTTGACTCTAT
TGTCTGGACCTTCAACACAACCCCTTTGTCACCATAAGCCAGAAGGGGGCACTATCATAGT
GACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAG
CAAACGTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCA
GCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCAT
GGGTCTGCAGAGCAATAAGAATGGCACCTGTGACCAATCTGACATGCTGCATGGAACATGG
GGAAGAGGGATGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGG
GTCCATCCTCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCCTGCCAG
GAACCCGTCAAGCAGAAACTCTCAAGCCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGC
TGATGACCCAGATTCCCTCATGGCCTCCTGTCTCCTGTTGGCTGCCCTCCTGCTCAGTCT
CTTGTACTGGGGCTATTCTTGGTTCTGAAGAGAGAGACAAGAAGAGTACATTGAAGA
GAAGAAGAGAGTGGACATTGTCGGAAACTCCTAACATATGCCCTATTCTGGAGAGAACAC
AGAGTACGACACAATCCCTCACACTAACAGAACATCCTAAAGGAAGATCCAGCAAATACGGT
TTACTCCACTGTGGAAATACCGAAAAAGATGGAAAATCCCCACTCACTGCTCACGATGCCAGA
CACACCAAGGCTATTGCCTATGAGAATGTTATCTAGACAGCAGTGCACTCCCCTAAGTCTCT
GCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAAASGPVKELVGSVGGAVTFPLSKVKQVDSIVWTFNTTPLVTI
QPEGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYE
HLSKPKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGES
DMTFICVARNPVSRNFSPIALKCEGAADDPDSSMVLLCLLLVPLLLSEVLGLFLWFLKR
ERQEEYIEEKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMEN
PHSLLTMPDTPRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAAT
GGATTCA~~G~~CCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATA~~C~~CTCTAATTGTC
AGCTTAGTTGAGGAAGACCAATTTCTAAAACCCC~~A~~TCTCTGCTTGAGTGGTGGT~~CC~~CA
GGAATTATAGGAGCAGGTCTGATGCCATTCCAGCAACAACAATGTC~~CT~~TGACAGCAAGAAAA
AGAGCGTGCTGCAACAA~~C~~AGAACAGAACTGGAATGTTCTTCATCATTTCAGTGTGATCACAGTC
ATTGGTGCTCTGTATTGCATGCTGATATCCATCCAGGCTCTTAAAAGGT~~C~~CTCTCATGTG
AATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTCATTGAAAAACATCAGTGACATT~~C~~
CCAGAAC~~T~~CTCAACTGCAGTGGTTTCAATGACTCTTGAC~~C~~CTC~~A~~TGGTTCAAT
AAACCCACCAGTAACGACACC~~A~~TGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCT
GAAGAAAAACAAACATAGGCTTATCCACTCTCAGTATTTAGGTCTATTGCTTGGAAATT
CTGGAGGT~~C~~CTGTTGGCTCAGTCAGATAGTCATCGGTT~~C~~GGCTGTGTGGAGTC
TCTAAC~~G~~GAAGAAGTCAAATTGTGTAGT~~T~~TAATGGGAATAAAATGTAAGTATCAGTAGTTGA
AAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLVLLLGVVNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAI
PATTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANC
EFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHF
SVFLGLLLGVILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGT
GAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGA
GCAGATCCGTGGGCTGCAGACCCCCGCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAA
CTGTGACATGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAAGC
CAATGACCCATTGCCAATAAGACGATCCCTCTACTATGACTGGAAAAACCTGCAGCTGAG
CGGACTGATCTGGGAGGGCTCCTGGCCATTGCTGGATCGCGCAGTCTGAGTGGCAAATG
CAAATACAAGAGCAGCCAGAACGACAGCACAGTCTGTACCTGAGAAGGCCATCCCACTCAC
TCCAGGCTCTGCCACTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAAC
ACTGGCCCCCAGCACCTCCTCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGG
GCAGGCTGTTAGGCCCTTCTGATCAGGAGGCTTCTTATGAATTAAACTGCCACCACC
CCCTCA

FIGURE 50

MERVTLALLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKY
KSSQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAAGGCTGCAGACCCAGAGGGAG
 GGAGGACAGGGAGTCGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGC
 AAGGAGGAGACCCTGGTGGAGGAAGACACTCTGGAGAGAGAGGGCTGGCAGAGATGAAG
 TTCCAGGGGCCCTGGCCTGCCTCCTGCTGGCCCTGCCTGGCAGTGGGAGGCTGGCC
 CTGCAGAGCGGAGAGGAAAGCACTGGACAAATATTGGGGAGGCCCTGGACATGGCCTGGGA
 GACGCCCTGAGCGAAGGGTGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCT
 AAAGTCAGTGAGGCCCTGGCCAAGGGACCAGAGAACAGCAGTTGGCACTGGAGTCAGGCAGGTT
 CCAGGCTTGGCGCAGCAGATGCTTGGCAACAGGGTCGGGAAGCAGCCCAGCTCTGGGA
 AACACTGGGCACGAGATTGGCAGACAGGCAGAACAGATGTCATTGACACGGAGCAGATGCTGTC
 CGCGGCTCCTGGCAGGGGTGCCTGGCACAGTGGTGCTTGGAAACTCTGGAGGTCTGGACT
 ATCTTGGCTCTCAAGGTGGCCTGGAGGCCAGGGCAATCCTGGAGGTCTGGACT
 CCGTGGGTCCACGGATAACCCGGAAACTCAGCAGGCAGCTTGGAAATGAATCCTCAGGGAGCT
 CCCTGGGGTCAAGGAGGCAATGGAGGCCACAAACTTGGGACCAACACTCAGGGAGCTGTG
 GCCCAGCCTGGCTATGGTCAGTGAGAGGCCAGCAACCAGAACATGAAGGGTGCACGAATCCCCA
 CCATCTGGCTCAGGTGGAGGCTCCAGCAACTCTGGGGAGGCAGCGGCTCACAGTCGGCAGC
 AGTGGCAGTGGCAGCAATGGTACAACAAACATGGCAGCAGCAGTGGTGGCAGCAGTGGCAGC
 AGCAGCAGTGGCAGCAGCAGCAGTGGCGCAGCAGTGGCAGCAGTGGCAGCAGTGGCAGC
 AGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGTCCTCCTGGGATCCAGCACCGGCTCCTCC
 TCCGGCAACCACGGTGGAGCGGGAGGAAATGGACATAAACCGGGTGTAAAAGCCAGGG
 AATGAAGCCCGGGAGCGGGAAATCTGGATTAGGGCTTCAGAGGACAGGGAGTTCCAGC
 AACATGAGGGAAATAAGCAAAGAGGGCAATGCCCTCCTGGAGGCTCTGGAGACAATTATCGG
 GGGCAAGGGTCGAGCTGGGCAGTGGAGGAGGTGACGCTGTTGGAGTCAATACTGTGAAC
 TCTGAGACGTCTCCTGGATTTAAGTACCTTGTGAAAGAATTAAATCCAAGCTG
 GGTTCATCAACTGGGATGCCATAAACAAAGGACCAAGAGAAGCTCTGCATCCCGTGACCTCCA
 GACAAGGGAGCCACCAGATTGGATGGAGGCCACACTCCCTCCTAAAACACCACCCCTCTCA
 TCACTAATCTCAGCCCTGCCCTGAAATAACCTTAGCTGCCCAACAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAA
GSKVSEALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGAD
AVRGSWQGVPGHSGAWETSGGHGIFGSQGGLGGQQGNPGGLGTPWVHGYPGNSAGSGFMNPQ
GAPWGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGSSNSGGSGSQS
GSSGSGSNGDNNNGSSGGSSSGSSGGSSGGSSGSGSNGSGSRGDSGSESSWGSSTG
SSSGNHGGSGGGNGHKPGCEKPGNEARGSGESGIQGFRQGVSSNMREISKEGNRLGGSDN
YRGQGSSWGSGGGDAVGGVNTVNSETSPGMNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70,
74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158,
155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205,
218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252,
249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281,
279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298,
295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328,
323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387,
383-389, 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGGACAAGCTGCTCCGACAGAAGG**ATG**TCGCTGCTGAGCCTGCCCT
 GGCTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCT
 GGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACGTGCCGCCGGCTCCAGT
 GTTTCCCACAGCCCCAAAACGGAACCTGGTTGGGTACCTGGCCTGATCACTCCTACAG
 AGGAGGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTACGGTATGGC
 TGGGTCCCATCATCCCCTTCATCGTTTATGCCACCCTGACACCATCCGGTCTATCACCAATG
 CCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGAG
 AAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGGCCACCGTCGGATGCTGACGCCGCCT
 TCCATTCAACATCCTGAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTG
 ACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTGAGCACATCAGCC
 TCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTGACAGCCATTGTCAGGAGAGGC
 CCAGTGAATATATTGCCACCCTTGGAGCTCAGTGCCCTTAGAGAAAAGAACGCCAGCATA
 TCCTCCAGCACATGGACTTCTGTATTACCTCTCCATGACGGCGGCGCTTCCACAGGCCT
 GCCGCCTGGTGCATGACTTCACAGACGCTGTCACTCCGGAGCGCGTCGCACCCTCCCCACTC
 AGGGTATTGATGATTTTCAAAGACAAAGCCAAGTCCAAGACTTGGATTGATGTGC
 TTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGATATAAGAGCAGAGGCTG
 ACACCTTCATGTTGGAGGCCATGACACCACGCCAGTGGCCTCTGGTCTGTACAACC
 TTGCGAGGCACCCAGAATACCAGGAGCGCTGCCACAGGAGGTGCAAGAGCTCTGAAGGACC
 GCGATCCTAAAGAGATTGAATGGGACGACCTGGCCAGCTGCCCTCCTGACCATGTGCGTGA
 AGGAGAGCCTGAGGTTACATCCCCAGCTCCCTCATCTCCGATGCTGCACCCAGGACATTG
 TTCTCCAGATGGCCAGTCATCCCCAAAGGCATTACCTGCCTCATCGATATTATAGGGTCC
 ATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACCCCTCCGCTTGACCCAGAGA
 ACAGCAAGGGAGGTACCTCTGGCTTTATTCCCTCCGCAGGGCCAGGAACGTGCATCG
 GGCAGGCAGTCGCCATGGCGAGATGAAAGTGGCCTGGCGTTGATGCTGCTGCACCTCCGGT
 TCCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAATTGATCATGCGCGCCAGGGCGGGC
 TTTGGCTGCCGGTGGAGCCCTGAATGTAGGCTTGCAG**TGA**CTTCTGACCCATCCACCTGTT
 TTTTGCAATTGTCATGAATAAAACGGTGCTGTCAAA

FIGURE 54

MSLLSLPWGLRPVAMSPWLLLLLVGSWLLARILAWTYAFYNNCRLQCFQPPKRNWFUGH
LGLITPTEEGLKDKSTQMSATYSQGFTVWLGPITIPFIVLCHPDTIRSITNASAAIAPKDNLFIR
FLKPWLGEGLLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRL
DMFEHISLMTLDSLQKCIFSFDHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSD
GRRFHACRLVHDFTDAVIREERRTLPTQGIDDEFKDKAKSKTLDFIGVLLSKDEDGKALSD
EDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCRQEVQELLKDRDPKEIEWDDLAQL
PFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHNPTVWPDPPEVYD
PFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLEL
IMRAEGGLWLRVEPLNVGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTCCTC**ATG**GGACCAGTGAAACAGCTGAAGCGAATGTT
GAGCCTACTCGTTGATTGCAACTATCATGGTGCTGTTGTGTTGCACTTACCCGTGTTCT
GCCTTTGGTGGCATAACAAGGGACTTGCACTTATCTCTGCATTTGCAGTCCTGGCATTG
ACGTGGTACAGCCTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTG
TGTCTTGCA**TAA**TTCATGCCAGTTTATGAAGCTTGGAGGCACATGGACAGAAGCTGGT
GGACAGTTTGTAACTATCTCGAACACCTCTGTCTTACAGACATGTGCCTTTATCTGCAGC
AATGTGTTGCTTGTGATTGAAACATTGAGGGTTACTTTGGAAGCAACAATACATTCTGAA
CCTGAATGTCAGTAGCACAGGATGAGAAGTGGTTCTGTATCTGTGGAGTGGAAATCTCCTC
ATGTACCTGTTCTCTGGATGTTGTCCTGAATTCCATGAATACAAACCTATTCA
AACAGCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGALIFCILQSLALTWYSLSFIPF
ARDAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGCAGTGGAGCACCCAGCAGGCCAAC**ATG**CTCTGTCTGTGC
 CTGTACGTGCCGGTCATCGGGGAAGCCCAGACCAGTTCACTTGTCAAGGAGCTTCACTTGTCAAGGAGCTCCCTGCCAGCTG
 AAGTCCATTTCAGCTCAGTGTCTCATCCCCCTCCAGGAATTCTCCACCTACCGCCAGTGGAAAGCAGAAAATT
 GTACAAGCTGGAGATAAGGACCTTGATGGCAGCTAGACTTGAGAAGAATTGTCCATTATCTCCAAGATCATGAG
 AAGAAGCTGAGGCTGGTTAAGATTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATGCAG
 TCCCTGCAGGGACTGGGAGTCAAGATATCTGAACAGCAGGCAGAAAAATTCTCAAGAGCATGGATAAAAACGGC
 ACGATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCCACCCGTGGAAAACATCCCCGAGATCATC
 CTCTACTGGAAGCATTCCACGATCTTGATGTGGGTGAGAATCTAACGGTCCCAGTGGTACAGTGGAGGAG
 AGGCAGACGGGATGTGGAGACACCTGGTGGCAGGAGGTGGGCAGGGCGTATCCAGAACCTGCACGGCC
 CCCCTGGACAGGCTCAAGGTGCTCATGCAGGTCCATGCCCTCCGCAGCAACACATGGCATCGTTGGTGGCTTC
 ACTCAGATGATTGAGAAGGGAGGGCCAGGTCACTCTGGCGGGCAATGGCATCAACGTCTCAAATTGCC
 GAATCAGCCATCAAATTATGGCCTATGAGCAGATCAAGCGCTTGTGGTAGTGACCAGGAGACTCTGAGGATT
 CACGAGAGGCTTGTGGCAGGGCCTTGGCAGGGCATCGCCAGAGCAGCATCTACCCAAATGGAGGTCTGAAG
 ACCCGGATGGCGCTGCGGAAGACAGGCCAGTACTCAGGAATGCTGGACTGCGCCAGGAGGATCTGGCAGAGAG
 GGGGTGGCCGCCTCTACAAAGGCTATGTCCCCAACATGCTGGGCATCATCCCTATGCCGCATCGACCTGCA
 GTCTACGAGACGCTCAAGAATGCCTGGCTGCAGCACTATGCAGTGAAACAGCGCGGACCCCGCGTGTGCTC
 CTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGTACCCCTGGCCCTAGTCAGGACCCGGATG
 CAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTCTTCAAACATATCCTGCGGACCGAG
 GGGGCCTCAGGGCTGTACAGGGGCTGGCCCCAACATTCACTGAAGGTCACTCCAGCTGTGAGCAGCTACGTG
 GTCTACGAGAACCTGAAGATCACCTGGCGTGCAGTCGG**TGA**CGGGGGAGGGCCGGCAGTGGACTCG
 CTGATCCTGGCCGCAGCCTGGGTGTGCAGCCATCTCATTCTGTGAATGTGCCAACACTAACGCTGTCTCGAGCC
 AAGCTGTAAAACCTAGACGCACCCGCAGGGAGGGTGGGGAGAGCTGGCAGGCCAGGGCTGTGCTGCTGACC
 CCAGCAGACCCCTCTTGTGGTCCAGCGAAGACCACAGGCATTCCCTAGGGTCCAGGGTCAGCAGGCTCCGGCT
 CACATGTGAAGGACAGGACATTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGGCTTAGTTC
 TTCCATTTCACCCCTGCAGCCAGTGTGGCCACGGCCCTGCCCTCTGGCTGCCGTGCATCTCCCTGTGCCCT
 CTTGCTGCCCTGCCTGTCTGAGGTAAGGTGGAGGGCTACAGCCACATCCCACCCCTGTCCAATCCC
 ATAATCCATGATGAAAGGTGAGGTACGTGGCCTCCAGGATGGCCCCACCTCAGAACCAAACACTGACGCCAATT
 GGCTGTGAAGGAAGAGGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGCTCTCG
 GCCATGCTGGAGTGCAGGGGCTCGGGCTGCCCTGGCTGCACAGAAGGCAAGTGTGGCTCATGGT
 CTCTGAGCTGGCCTGGACCCCTGTCAAGGATGGCCCCACCTCAGAACCAAACACTGTCCTGAGCTGGCATGAG
 GGCAGTGGAGCACCATTGAGGGCAAGGGCAGAGCTTGTGTTCTGGGGAGGGAGGAAAGGTTGG
 AGGCCTTAATTATGGACTGTTGGAAAAGGGTTTGTCCAGAACGCGAACAAATGAGCGACTTGTG
 TTCCAGAGGAAGACGAGGGAGCAGGAGCTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGTTCTGTC
 CAACCCAGCAGGGCGCAGCGGGACCAGCCCCACATTCCACTTGTGTCAGTGTGGAACCTATTATTGTA
 TTTATTGAACAGAGTTATGTCCTAACTATTATAGATTGTTAATTAAATAGCTGTGCAATTTCAGTTCAT
 TTTTATTGATATTGTCATGGTGTGATTGACTCTCCAAAGCCGCCAGTGGATGGAGGGAGGAGGAGAA
 GGGGGCCTGGGCCGCTGCAGTCACATCTGTCCAGAGAAATTCTTTGGACTGGAGGCAGAAAAGCGGCCAG
 AAGGCAGCAGCCCTGGCTCCTTGGCAGGTTGGGAAGGGCTTGGCCCTAGGATTTCAGGGTT
 GACTGGGGCGTGGAGAGAGAGAGGGAGGAACCTCAATAACCTGAGGTTGGAATCCAGTTATTCTGCGCTGCG
 GGGTTCTTATTCACTCTTGTGAATGTCAAGGCAGTGAGGTGCCTCTCACTGTGAATTGTGGTGGCGGG
 GGCTGGAGGAGAGGGTGGGGCTGGCTCCGCTCCAGCCTCTGCTGCCCTGCTTAACAAATGCCGGCAA
 CTGGCGACCTCACGGTTGCACTCCATTCCACCAAGAACCTGATGAGGAAATCTCAATAGGATGCAAAGATC
 AATGCAAAATTGTTATATGAACATATAACTGGAGTCGTCAAAAGCAAATTAAAGAAAGAATTGGACGTTAGA
 AGTTGTCAATTAAAGCAGCCTCTAATAAAAGTTGTTCAAAGCTGAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSYRQWKQKIVQAGDKDL
DGQLDFEEFVHYLQDHEKKRLVFKILDKKNDGRIDAQEIMQLSLRDLGVKISEQQAEKILKSM
DKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKSTIFDVGENLTVPDEFTVEERQTGMWWRH
LVAGGGAGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLK
IAPESAIKFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRKTGQ
YSGMLDCARRILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLNAWLQHYAVNSADPGVF
VLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAP
NFMKVIPAVSISYVVYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGCAGCTCCACTCAGCCAGTACCCAGATA CGCTGGAACCTTCCCCAGCC**ATG**
 GCTTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCA
 ATTGCACTCATCATTGGCTTGGTATTCAGGGAGACACTCCATCACAGTCACACTGTGCC
 TCAGCTGGAACATTGGGAGGATGGAATCCTGAGCTGCACTTGAACCTGACATCAAACCT
 TCTGATATCGTATAACAATGGCTGAAGGAAGGTGTTAGGCTTGGCCATGAGTTCAAAGAA
 GGCAAAGATGAGCTGTCGGAGCAGGATGAAATGTTAGGCGGACAGCAGTGTGATG
 CAAGTGATAGTTGGCAATGCCCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACC
 TACAAATGTTATATCATCACTCTAAAGGCAAGGGAAATGCTAACCTGAGTATAAAACTGGA
 GCCTTCAGCATGCCGGAAAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCAGGTGTGAG
 GCTCCCCGATGGTCCCCCAGCCCACAGTGGCTGGCATCCAAAGTTGACCAGGGAGCCAAC
 TTCTCGGAAGTCTCCAATACCAGCTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTG
 TCTGTGCTCTACAATGTTACGATCAACACACATACTCCTGTATGATTGAAAATGACATTGCC
 AAAGCAACAGGGATATCAAAGTGACAGAATCGGAGATCAAAAGGCGGAGTCACCTACAGCTG
 CTAAACTCAAAGGCTCTGTGTCTCTTCTTGCATCAGCTGGCACTTGCCT
 CTCAGCCCTTACCTGATGCTAAAAT**TAAT**TGTCCTGGCCACAAAAAGCATGCAAAGTCATTG
 TTACAACAGGGATCTACAGAACTATTCACCACCAAGATATGACCTAGTTATATTCTGGGA
 GGAAATGAATTCATATCTAGAAGTCTGGAGTGAGCAAACAAGAGCAAGAAACAAAAAGAAGCC
 AAAAGCAGAAGGCTCCAATATGAACAAAGATAATCTATCTCAAAGACATATTAGAAGTTGGG
 AAAATAATTCATGTGAACTAGACAAGTGTGTTAAGAGTGATAAGTAAAATGCACGTGGAGACA
 AGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGTCACCTGGGAGTGAGAGGACAGGAT
 AGTGCATGTTCTTGTCTGAAATTTAGTTATATGTGCTGTAATGTTGCTCTGAGGAAGCC
 CCTGGAAAGTCTATCCAACATATCCACATCTTATATTCCACAAATTAGCTGTAGTATGTAC
 CCTAAGACGCTGCTAATTGACTGCCACTCGCAACTCAGGGCGGCTGCATTTAGTAATGGG
 TCAAATGATTCACTTTATGATGCTTCAAAGGTGCCTGGCTCTTCCAACTGACAAA
 TGCCAAAGTTGAGAAAAATGATCATATAATTAGCATAAACAGAGCAGTCGGGGACACCGATTT
 TATAAAATAAACTGAGCACCTTCTTTAAACAAAAA
 AAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIK
LSDIVIQWLKEGVGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAG
TYKCYIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVDQGA
NFSEVSNTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRRSHLQ
LLNSKASLCVSSFFAISWALLPLSPYIMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTAGAATCACCA**ATG**GCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACA
 AGCACCAGGAGCCCCTCCGGGTAGCTACTACCCTGGACCCCCAATAGTGGAGGGCAGTATGG
 TAGTGGCTACCCCTGGTGGTTATGGGGCTGCCCCTGGAGGGCCTATGGACCC
 AGCTGGTGGAGGGCCCTATGGACACCCAATCCTGGATGTTCCCTCTGGAACCTCCAGGAGG
 ACCATATGGCGGTGCAGCTCCGGGGCCCTATGGTCAGCCACCTCCAAGTTCTACGGTGC
 CCAGCAGCCTGGGCTTATGGACAGGGTGGCGCCCTCCCAATGTGGATCCTGAGGCCTACTC
 CTGGTCCAGTCGGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGC
 CCTGGTCAACTGCAATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTT
 TGACAAGACCAAGTCAGGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCC
 GCAGTGGAAAGAACCTCTCCAGCAGTATGACCGGGACCGCTCGGCTCCATTAGCTACACAGA
 GCTGCAGCAAGCTCTGTCCAAATGGCTACAACCTGAGCCCCAGTTCACCCAGCTCTGGT
 CTCCCGCTACTGCCACGCTCTGCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTG
 CACCCAGCTGCAGGTGCTGACAGAGGCCTCCGGGAGAAGGACACAGCTGTACAAGGCAACAT
 CCGGCTCAGCTCGAGGACTTCGTACCATGACAGCTCTCGGATGCT**TGA**CCCAACCCT
 GTGGAGAGTGGAGTGCACCAGGACCTTCCTGGCTTCTAGAGTGAGAGAAGTATGTGGACA
 TCTCTTCTTCCTGTCCCTCTAGAAGAACATTCTCCCTGCTTGATGCAACACTGTTCCAAA
 AGAGGGTGGAGAGTCCTGCATCATGCCACCAAATAGTGAGGACCGGGCTGAGGCCACACAG
 ATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGAATGTCCTGATGCCATGAGCAGTTGAGTG
 GCACAGCCTGGCACCAGGAGCAGGTCTGTAATGGAGTTAGTGTCCAGTCAGCTGAGCTCCA
 CCCTGATGCCAGTGGTGAGTGTTCATCGGCCTGTTACCGTTAGTACCTGTGTTCCCTCACCAG
 GCCATCCTGTCAAACGAGCCCATTCTCCAAAGTGGAAATCTGACCAAGCATGAGAGAGATCT
 GTCTATGGGACCAGTGGCTGGATTCTGCCACACCCATAAATCCTGTGTTAACTTCTAGC
 TGCCTGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGCATCTTGGCCAGGCTCTGCC
 CCCTGCAGCTGGGACCCCTCACTGCCTGCCATGCTCTGCTCGGCTTCAGTCTCAGGAGACA
 GTGGTCACCTCTCCCTGCCAATACTTTTAATTGCATTTCATTGGGCCAAAG
 TCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGP
YGHPNPGMFPSTPGGPYGGAAPGGPYGQPPPSSYGAQQPGLYQGGAPPNVDPEAYSWFQSV
DSDHSGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTSGRIDVYGFSAWKFIQQWKNL
FQQYDRDRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQV
LTEAFREKDTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTC
TCTCCCAGGCCTCTTGCCCCGGAGCATCGGTGTTGGAGGAGAAAGTTCCAAAACCTCGG
GACCAACTTGCCCTCAGCTCGGACAACCTCCTCCACTGGCCCTCTAACTCTGAACATCCGCA
GCCCGCTCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCCCTGAAGCTCAGCGTGCCTCC
ATCAGATGGCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGCCCTCATCGTGGGGCT
GCCTGCCATGGATT CCTGGCCCCCTGAGGATCCTGGCAGATGATGGCTGCTGCGGCTGAGGA
CCGCCTGGGGGAAGCGCTGCCTGAAGAACTCTTACCTCTCCAGTGCTGCAGGCCCTCGCTCC
GGGCAGTGGCCCTTGCCTGGGGAGTCTTCTCCGATGCCACAGGCCTCTCACCTGAGGCTTC
ACTCCTCCACCAGGACTCGGAGTCCAGACGACTGCCCGTTCTAATTCACTGGGAGGCCGGGG
AAAAATCCTTCCAACGCCCTCCCTGGTCTCATCCACAGGGTTCTGCCTGATCACCCCTG
GGGTACCTGAATCCCAGTGTGCTGGGGAGGTGGAGGCCCTGGACTGGTTGGGAACGAG
GCCCATGCCACACCCTGAGGGAATCTGGGTATCAATAATCAACCCCCAGGTACAGCTGGGG
AAATATTAATCGGTATCCAGGAGGCAGCTGGGAAATATTAATCGGTATCCAGGAGGCAGCTG
GGGAATATTAATCGGTATCCAGGAGGCAGCTGGGAAATATTCAATACCCAGGTATCAA
TAACCCATTCCTCCGGAGTTCTCCGCCCTGGCTCTGGCTTTGGAACATCCCAGCTGGCTT
CCCTAACCTCCAAGCCCTAGGTTGCAGTGGGCTAGAGCACGATAGAGGGAAACCCAACATT
GGGAGTTAGAGTCCTGCTCCGCCCTTGCTGTGGCTCAATCCAGGCCCTGTTAACATGT
TTCCAGCACTATCCCCACTTTCACTGCCTCCCTGCTCATCTCCAATAAAATAAAAGCACTT
ATGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 64

MQGRVAGSCAPLGLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQP
ALDPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDR
LGEALPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGK
ILSQRPPWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRMPHPEGIWGGINNQPPGTWGN
INRYPGGSWGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPVGVLRPPGSSWNIPAGFP
NPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGG
TCTGGGCTGCCCTTGTCCCTCTTGACCCCTCCTGGCAGCTCACATGGAACAGGGCCGGGT
ATGACTTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATTCCCTATGAGTCCAGCTTC
CTGGAATTGCTTGAAAAGCTCTGCCTCCTCCATCTCCCTCAGGGACCAGCGTCACCCTC
CACCATGCAAGATCTAACACCATGTTGTCTGCAACACATGACAGCCATTGAAGCCTGTGTCC
TTCTTGGCCGGGCTTTGGGCCGGGATGCAGGAGGCAGGCCCGACCCTGTCTTCAGCAG
CCCCCACCCTCCTGAGTGGCAATAAATAAATTGGTATGCTG

FIGURE 66

MGSGLPLVLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFELLEKLCLLHLPSGTSV
TLHHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCGG
GCCAGGTGCCCGTCGCAGGTGCCCTGGCCGGAGATCGGGTAGGAGGGCGAGCGCGAGAAG
CCCCTCCTCGCGCTGCCAACCCGCCACCCAGCCC**ATG**GCGAACCCGGCTGGGCTGCTT
CTGGCGCTGGCCTGCCGTTCTGCTGGCCCGCTGGGCCGAGCCTGGGGCAAATACAGACC
ACTTCTGCAAATGAGAATAGCACTGTTGCCTCATCCACCAGCTCCAGCTCCGATGGCAAC
CTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTTTGGCTGCCTGCTCCTG
GCTGTGGGCTGGCACTGTTGGTGCAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTAC
CGGCCAGTAGCGAGGAGCAGTTCTCCATGCAGCCGAGGCCCGGGCCCTCAGGACTCCAAG
GAGACGGTGCAGGGCTGCCATC**TAG**GTCCCCTCTGCATCTGTCTCCCTTCATTGC
TGTGTGACCTTGGGAAAGGCAGTGCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAG
CAGGGAAGAAGGTACTTCAAAGACTCTGCCCTGAGGTCAAGAGAGGGATGGGCTATTCACTT
TTATATATTATATAAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLARWGRAWGQIQTTSANENSTVLPSSSTSSSDGNLRPEAITAIIVV
FSLLAALLLAVGLALLVRKLREKRQTEGYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACA**ATGGGGTTATT**CAGAGGTTTGTTCCTCTAGTTCTGTGCCTGCTGCAC
 CAGTCAAATACTCCTCATTAAGCTGAATAATAATGGCTTGAAAGATATTGTCATTGTTATAGATCCTAGTGTG
 CCAGAAGATGAAAAAATAATTGAACAAATAGAGGATATGGTACTACAGCTTACGTACCTGTTGAAGCCACA
 GAAAAAAAGATTTTTCAAAAATGTATCTATATTAAATTCTGAGAATTGGAAGGAAATCCTCAGTACAAAGG
 CCAAAACATGAAAACCATAAACATGCTGATGTTATAGTGCACCACTACACTCCCAGGTAGAGATGAACCATAAC
 ACCAAGCAGTTCACAGAATGTGGAGAGAAAGGCGAATACATTCACTCACCCCTGACCTCTACTTGGAAAAAAA
 CAAAATGAATATGGACCACCAGGCAAACGTGTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGATGAG
 TACAATGAAGATCAGCCTTCTACCGTCTAAGTCAGTCAAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCT
 GGTAGAAATAGAGTTATAAGTGTCAAGGAGGAGCTGTTCTGAGTAGAGCATGCAGAATTGATTCTACAAACAAAA
 CTGTATGGAAAAGATTGTCAATTCTTCCTGATAAAAGTACAAACAGAAAAGCATCCATAATGTTATGCAAAGT
 ATTGATTCTGTTGTAACGAAAAACCCATAATCAAGAAGCTCCAAGCCTACAAACATAAAGTGC
 AATTAGAAGTACATGGGAGGTGATTAGCAATTCTGAGGATTAAAACACCATACCCATGGTACACCCACCT
 CCTCCACCTGTCTCTCATTGCTGAAGATCAGTCAAAGAATTGTTGCTTAGTTCTGATAAGTCTGGAAGCAGTGC
 GGGGTAAGGACGCCTAAATCGAATCAAGCAGCAAACATTCTGCTGCAGACTGTTGAAAATGGATCC
 TGGGTGGGGATGGTTCACTTGATAGTACTGCCACTATTGTAATAAGCTAATCCAATAAAAGCAGTGATGAA
 AGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGAACCTCCATCTGCTCTGGAATTAAATATGCA
 TTTCAGGTGATTGGAGAGCTACATTCCAACTCGATGGATCCGAAGTACTGCTGACTGATGGGAGGATAAC
 ACTGCAAGTTCTGTATTGATGAAGTGAACAAAGTGGGCCATTGTCATTGTTATTGCTTGGGAAGAGCTGCT
 GATGAAGCAGTAATAGAGATGAGCAAGATAACAGGGAGGAAGTCATTGTTATGTTCAGATGAAGCTCAGAACAAAT
 GGCCTCATTGATGCTTGGGCTCTACATCAGGAAATACTGATCTCTCCAGAAGTCCCTCAGCTCGAAAGT
 AAGGGATTAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAATTGATAGTACAGTGGGAAAGGACACG
 TTCTTCTCATCACATGAAACAGTCTGCCTCCAGTATTCTCTGGATCCCAGTGGAAACAATAATGGAAAAT
 TTCACAGTGGATGCAACTTCAAATGGCCTATCTCAGTATTCCAGGAAC TGCAAAGGTGGCACTTGGCATAAC
 AATCTCAAGCCAAGCGAACCCAGAAACATTAACATTACAGTAACCTCTCGAGCAGCAAATTCTCTGTCCT
 CCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTCCCCAGCCAAATGATTGTTACGCAGAAATT
 CTACAAGGATATGTACCTGTTCTGGAGCCAATGTGACTGCTTCAAGAATGATGGAGTCTACTCCAGGTATTTACAGCA
 TTGGAACCTTGGATAATGGTGCAGGCCTGATTCTCAAGAATGATGGAGTCTACTCCAGGTATTTACAGCA
 TATACAGAAAATGGCAGATATAGCTTAAAGTCTGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAAATTACGG
 CCTCCACTGAATAGAGCCGCGTACATACCAGGCTGGTAGTGAACGGGAAATTGAAGCAAACCCGCAAGACCT
 GAAATTGATGAGGATACTCAGACCACCTGGAGGATTCAGCCGAACAGCATCCGGAGGTGCATTGTTGATCA
 CAAGTCCAAGCCTCCCTGCCTGACCAATACCCACCAAGTCAAATCACAGACCTGATGCCACAGTTCATGAG
 GATAAGATTATTCTTACATGGACAGCACCAGGAGATAATTGATGTTGAAAAGTTCAACGTTATATCATAAGA
 ATAAGTGCAAGTATTGATCTAAGAGACAGTTGATGATGCTCTCAAGTAAATACTACTGATCTGTCACCA
 AAGGAGGCCAAGCCTTGCATTAAACAGAAAATCTCAGAAGAAAATGCAACCCACATATT
 ATTGCCATTAAAAGTATAGATAAAAGCAATTGACATCAAAGTATCCAACATTGCAACAGTAACATTGTTATC
 CCTCAAGCAAATCCTGATGACATTGATCCTACACCTACTCCTACTCCTACTCCTGATAAAAGTCATAAT
 TCTGGAGTTAATATTCTACGCTGGATTGTCATTGGCTGTTGTAATTGTTAACTTATTAAAGTACC
 ACCATT**TGA**ACCTTAACGAAGAAAAAAATCTCAAGTAGACCTAGAAGAGGTTAAAAACAAAACAATGTAAG
 TAAAGGATATTCTGAATCTTAAATTGATCCCAGTGTGATCATAAACTCATAAAAATAATTAAAGATGTCG
 GAAAAGGATACTTGTATAAAACACTCATGGATATGTAAGGAAACTGTCAAGGATTTAAATAGTTCA
 TTTATTGTTATTGTAAGAAATAGTGTGAAACAAAGATCCTTTCTACTGATACCTGGTTGTATATT
 ATTTGATGCAACAGTTCTGAAATGATATTCAAATTGCAAGAAATTAAATCATCTATCTGAGTAGTCAA
 AATACAAGTAAAGGAGAGCAAAATAACACATTGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA
 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA

FIGURE 70

MGLFRGFVFLVLCLLHQSNNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTTASTYL
 FEATEKRFKKNVSLIPENWKENPQYKRPKHENKHADVIVAPPTLPGRDEPYTKQFTECGE
 KGEYIHFTPDLLGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCSA
 GISGRNRVYKCQGGSCLSACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVVEFCNE
 KTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVSLLKISQRIVCLVLDKSG
 SMGGKDRLNRMNQAACKHFLQLQTENGWSWGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTY
 PLGGTSICSGIKYAFQVIGELHSQLDGSEVLLTDGEDNTASSCIDEVKQSGAIVHFIALGRA
 ADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGTLNSNAWMN
 DTVIIDSTVGKDTFFLITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAY
 NLQAKANPETLTITVTSRAANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVT
 AFIESQNGHTEVLELLDNGAGADSFKNNDGVYSRYFTAYTENGRYSLKVRAGGANTARLKLRP
 PLNRAAYIPGWVVNGEIEANPPREIDEDTQTTLEDFSRTASGGAFVVSQVPSLPLPDQYPPS
 QITDLDATVHEDKIILTWTAPGDNFDVGKVQRYIIRISASILDRLDSFDDALQVNTTDLSPKE
 ANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPT
 PTPTPDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACTGCCCGGGCAGGGG
 TGACAACAGGTGTCATCTTTGATCTCGTGTGGCTGCCTCCTATTCAAGGAAAGACGCCAAGGTAATTT
 GACCCAGAGGAGCAATGATGTAGCCACCTCCTAACCTCCCTCTGAACCCCCAGTTATGCCAGGATTACTAG
 AGAGTGTCAACTCAACCAGCAAGCGGCTCCTCGGCTTAACCTGTGGTGGAGGAGAACCTTGTGGGGCTGC
 GTTCTCTTAGCAGTGCTCAGAAGTGACTTGCTGAGGGTGGACCAGAAGAAAGGAAAGGTTCCCTCTGCTGTTG
 GCTGCACATCAGGAAGGCTGTGATGGGAATGAAGGTGAAAACCTGGAGATTCACTTCAGTCATTGCTCTGCCT
 GCAAGATCATCCTTAAAAGTAGAGAAGAGCTGCTCTGTGTGGTGGTTAACCCAAGAGGAGAACACTCGTTCTAGAA
 GGAAATGGATGCAAGCAGCTCCGGGGCCCCAAACGCATGCTCCTGTGGCTAGCCCAGGGAAAGCCCTCCGTG
 GGGGCCCGGCTTGAGGGATGCCACCCTGCTGGACGCATGGCTGATTCTGAATGATGATGGTTCGCCGGGG
 CTGCTTGCCTGGATTCCCGGGTGGTTGCTGGTGCCTCTGCTGTGCTATCTCTGTCTGACATGTTG
 GCCTGCACCCAAAAGGTGACGAGGAGCAGCTGGCACTGCCAGGGCAACAGCCCCACGGGAAGGGAGGGTAC
 CAGGCCGTCTTCAGGAGTGGAGGGAGCAGCACCGCAACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTC
 AAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGGAATGGCAGTACCAAGCCAGCGATGCTGGCCTGGT
 CTGGACAGGAGCCCCCAGAGAAAACCCAGGCCACCTCTGGCCTCTGCACTCAGGTGGACAAGGCAGAG
 GTGAATGCTGGCGTCAAGCTGCCACAGAGTATGCAGCAGTGCCTTCGATAGCTTACTCTACAGAAGGTGTAC
 CAGCTGGAGACTGGCTACCCGCCACCCGAGGAGAACGCTGTGAGGAAGGACAAGCGGGATGAGTTGGTGGAA
 GCCATTGAATCAGCCTTGGAGACCCCTGAACAATCCTGCAGAGAACAGCCCCAATCACCGTCTTACACGGCCTCT
 GATTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTATGAGCTCACCTCAAAGGGGACCAC
 AAACACGAATTCAAACGGCTCATCTTATTCGACCATTAGCCCATCATGAAAGTGAAAATGAAAAGCTCAAC
 ATGGCCAACACGCTTATCAATGTTATCGCCTCTAGCAAAAGGGTGGACAAGTCCGGCAGTTCATGCAGAAT
 TTCAGGGAGATGTCATTGAGCAGGATGGGAGAGTCCATCTCAGTGTGTTACTTGGAAAGAAGAAATAAAT
 GAAGTCAAAGGAATACTGAAAACACTTCAAAGCTGCCACTTCAGGAACCTTACCTCATCCAGCTGAATGGA
 GAATTTCTCGGGAAAGGGACTTGATGTTGGAGCCGCTCTGGAAGGGAAGCAACGTCTCTCTTTCTGT
 GATGTGGACATCTACCTCACATCTGAATTCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAAGAAGGTATT
 TATCCAGTTCTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTGGAACAG
 CAGCTGGTCATAAAGAAGGAAACTGGATTGGAGAGACTTGGATTGGATGACGTGTAGTACCGGTAGAC
 TTCATCAATATAGGTGGTTGATCTGGACATCAAAGGCTGGGGGGAGAGGATGTGCACCTTATCGCAAGTAT
 CTCCACAGCAACCTCATAGTGGTACGGACGCCTGTGCGAGGACTCTCCACCTCTGGCATGAGAAGCGCTGCATG
 GACGAGCTGACCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCCACGGCCAGCTG
 GGCATGCTGGTGGTCAAGGCACGAGATAGAGGCTCACCTCGCAAACAGAAACAGAAGACAAGTAGCAAAAAAAC
TGAACTCCCAGAGAAGGATTGTTGGAGACACTTTCTTCCTTGCAATTACTGAAAGTGGCTGCAACAGAGA
 AAAGACTTCCATAAAGGACGACAAAAGAATTGGACTGATGGGTGAGAGATGAGAAAGCCTCGATTCTCTCTGT
 TGGCTTTTACAACAGAAATCAAATCTCGCTTGCCTGCAAAGTAACCCAGTTGCACCCCTGTGAAGTGTCT
 GACAAAGGCAGAATGCTGTGAGATTATAAGCCTAATGGTGTGGAGGTTTGATGGTGTGTTACAATACACTGAGA
 CCTGTTGTTGTGCTCATTGAAATATTGATTTAAGAGCAGTTGTAAGGAAATTCTATTGATGAAAG
 CAAGCATATTCTCCTCATATGAATGAGCCTATCAGCAGGGCTCTAGTTCTAGGAATGCTAAATATCAGAAGG
 CAGGAGAGGAGATAGGCTTATTGATGAGTACTGAGTACATTAAGTAAAATAAGGACCAAGAAAAGAAAAGAA
 ACCATAAATATCGTGTATTTCCCCAAGGATTAACCAAAAATAATCTGTTATCTTTGGTTGTCTTTAA
 CTGTCCTCGTTTTCTTTATTGAAATGCACTTTCTGAGTTGAGTTAGTCTGCTTATTGAGATGCA
 CACTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTACATTGATTTATATTGAGATGCA
 TTATGAGAACTTCAGTCAGTCAGGACATGAGACATAGGGAAAGGAATGGTTGAGTTAGTCTGCTTACAGAAAAGGAAACT
 CATTGAGACTGGTGTGATGAGTACCTAAAAGTCAGAAACCACATTCTCCTCAGAAGTAGGGACCGCTT
 CTTACCTGTTAAATAACCAAAGTATACCGTGTGAACCAAACAATCTCTTCAAAACAGGGTGTCTCCTCTGG
 CTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATTGTGAAAGATCAATCCATCTG
 CCAGAATCTAGTGGATGGAAGTTTGCTACATGTTATCCACCCAGGCCAGGTGGAAAGTAAGTGAATTATTT
 TAAATTAAAGCAGTTACTCAATCACCAAGATGCTCTGAAAATTGCAATTATTACCAACTATT
 TAAAATAAAATACAGTTAACATAGAGTGGTTCTCATTGAGTAAAGTCAAGGACCATGAG
 CTAATTATCTCTTGAGTCCTGCTCTGCTCACAGTAAACTCATGTTAAAAGCTTCAAGAACATTCAA
 GCTGTTGGTGTGTTAAAATGCATTGAGTTGACTGGTAGTTATGAAATTAAATTAAACACAGGCCA
 TGAATGGAAGGTGGTATTGCACAGCTAACAAATGATTGTGGATATGAA

FIGURE 72

MMMVRRGLLAWISRVVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQE
WEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHS
QVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIESALE
TLNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVK
NEKLNMANTLINVIVPLAKRVDKFRQFMQNREMCEQDGRVHLTVVYFGKEEINEVKGILEN
TSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPG
KKVFYPVLFSQYNPGIIYGHDAVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDL
DIKGWGGEDVHLYRKYLHSNLIVVRTPVRLFHLWHEKRCMDELTPEQYKMCMQSKAMNEASH
GQLGMLVFRHEIEAHLRKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGAGGGCAAAGAGGCAGCAAGAGAGTTGTCCTGGGGATC
 CAGAAACCCATGATAACCCTACTGAACACCGAATCCCCTGGAAGGCCACAGAGACAGAGACAGC
 AAGAGAACGAGAGATAAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCA
 CTCCTCCCTCCCTCTCTGCCTGTCCTAGTCCTCTAGTCCTCAAATTCCCAGTCCCTGC
 ACCCCTTCCTGGGACACTATGTGTTCTCCGCCCTCCTGCTGGAGGTGATTGGATCCTGGCT
 GCAGATGGGGTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATTGCCAGCCT
 TACCCCTGAGTGTGGAAACAATGCCAGTCGCCATCGATATTGAGACAGACAGTGTGACATT
 GACCCTGATTGCCTGCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGAC
 CTGCACAACAATGCCACACAGTCAACTCTCTGCCCTTACCCGTATCTGGGTGGACTT
 CCCCAGAAAATATGTAGCTGCCAGCTCCACCTGCACTGGGTCAAGAAAGGATCCCCAGGGGG
 TCAGAACACCAGATCAACAGTGAAGCCACATTGAGCTCCACATTGTACATTGACTCT
 GATTCCATGACAGCTTGAGTGAGGCTGCTGAGAGGCCTCAGGGCTGGCTGGCATC
 CTAATTGAGGTGGGTGAGACTAAGAATATAGCTTATGAACACATTGAGTCATTGCATGAA
 GTCAGGCATAAAGATCAGAACGACCTCAGTGCCTCCCTAACCTAACAGAGAGCTGCTCCCCAAA
 CAGCTGGGCAGTACTCCGCTACAATGGCTCGCTACAACCTCCCTGCTACCAGAGTGT
 CTGGACAGTTTATAGAAGGTCCCAGATTCAATGGAACAGCTGGAAAAGCTTCAGGG
 ACATTGTTCTCCACAGAACAGAGGAGCCCTTAAGCTTCTGGTACAGAAACTACCGAGGCCCTCAG
 CCTCTCAATCAGCGCATGGCTTGCCTTCAAGCAGGATCCTCGTATACCACAGGT
 GAAATGCTGAGTCTAGGTGTAGGAATCTGGTGGCTGTCTGCCTCTGGCTGTTAT
 TTCATTGCTAGAAAGATTGGAAGAACAGAGGCTGGAAAAGCGAAAGAGTGTGGTCTCACCTCA
 GCACAAGCCACGACTGAGGCATAAATTCCCTCTCAGATACCATGGATGTGGATGACTCCCTT
 CATGCCTATCAGGAAGCCTCTAAATGGGGTAGGATCTGCCAGAACACTGTAGGAGTAG
 TAAGCAGATGTCCTCCTCCCTGGACATCTTAGAGAGGAATGGACCCAGGCTGTCATTCC
 AGGAAGAACACTGCAGAGCCTCAGCCTCTCAAACATGTAGGAGGAATGAGGAAATCGCTGTG
 TTGTTAATGCAGAGANCAAACACTGTTAGTGCAGGGAGTTGGATATACCCCAAAGTC
 CTCTACCCCTCACTTTATGCCCTTCCCTAGATATACTGCAGGGATCTCTCCTTAGGATAA
 AGAGTTGCTGTTGAAGTTGTATATTTGATCAATATATTGGAAATTAAAGTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECGNNAQSPIDIQTDSVTFDPLPA
LQPHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQIN
SEATFAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQ
KTSVPPFNLRELLPKQLGQYFRYNGSLTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTE
EEPSKLLVQNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKI
RKKRLENRKSVVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCCTGGGGACGGGCAGTCCCTGTG
 TCTCTGGTGGTTGCCTAACACCTGCAAACATCACCTTCTTATCCATCAACATGAAGA**ATGTCC**
 TACAATGGACTCCACCAGAGGGTCTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTCA
 TCACAAATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTGTTGTCCTG
 ACAGCTCCAGAGAAGTGAAGAGAAAATCCAGAAGACCTTCCTGTTCCATGCAACAAATATA
 TCCAATCTGAAGTATAACGTCTGTGAATACTAAATCAAACAGAACGTGGTCCCAGTGT
 GTGACCAACCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTACTGCGTACACGT
 GAGTCCTCGTCCCAGGGCCCCCTGCCGTGCTCAGCCTCTGAGAAGCAGTGTGCCAGGACT
 TTGAAAGATCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTGCCATATCT
 ATTACCGTGTTCCTTTCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAA
 GAGAACACCCAGCAAATTGATTGATTGAAATGAATTGACAAAAGATTCTTGTG
 CCTGCTGAAAAAAATCGTATTAACTTTATCACCCCTCAATATCTGGATGATTCTAAAATTCT
 CATCAGGATATGAGTTACTGGAAAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCC
 AGCGGGAACCTGAGGCCCTCAGGAGGAAGAGGGAGGTGAAACATTAGGGTATGCTTCGCAT
 TTGATGGAAATTTGTGACTCTGAAGAAAACACGGAAGGTACTTCTCTCACCCAGCAAGAG
 TCCCTCAGCAGAACAAATACCCCGATAAAACAGTCATTGAATATGAATATGATGTCAGAAC
 ACTGACATTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTGCAGGAGGAGGTGTCCACACAA
 GGAACATTATTGGAGTCGCAGGCAGCGTTGGCAGTCTGGGCCAAACGTTACAGTACTCA
 TACACCCCTCAGCTCCAAGACTTAGACCCCTGGCCAGGAGCACACAGACTCGGAGGAGGG
 CGGAGGAAGAGCCATCGACGACCCCTGGTCAGGGATCCCCAAACTGGCAGGCTGTATT
 CCTTCGCTGTCCAGCTCGACCAGGATTCAAGGGCTGCGAGCCTCTGAGGGGGATGGGCTC
 GGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGGCCGGCTCCAGACAGGCCACCAGGAGAA
 AATGAAACCTATCTCATGCAATTGAGGAATGGGGTTATATGTCAGATGGAAA**TGA**
 TGCCAAACACTCCTTTGCCTTGTGCAAACAAAGTGAAGTCACCCCTTGATCCCA
 GCCATAAAAGTACCTGGGATGAAAGAAGTTTTCCAGTTGTCAGTGTCTGTGAGAATTACTT
 ATTTCTTCTATTCTCATAGCACGTGTGATTGGTCATGCATGTAGGTCTTTAACAA
 TGATGGTGGCCTCTGGAGTCCAGGGCTGGCCGGTTGTTATGCAGAGAAAGCAGTCAATA
 AATGTTGCCAGACTGGTGCAGAATTATTGAGGTGGGTGT

FIGURE 76

MSYNGLHQRFKELKLLTLCISISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQ
IYSNLKYNVSVLNTKSRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCA
RTLKDQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILIGNEFDKRF
FVPAEKIVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEVKHLGYA
SHLMEIFCDSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLOEEVS
TQGTLLESQAALAVLGPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRL
CIPSLSSFDQDSEGCEPSEGGLGEEGLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMCN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTTGAC
 ACCTGGGAAG**ATG**GCCGGCCGTGGACCTTCACCCTCTGTGGTTGCTGGCAGCCACCTT
 GATCCAAGCCACCCTCAGTCCCCTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAA
 GCTGACACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAG
 TGCCATGCGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTAACACCCGTCT
 GAAGCACATCATCTGGCTGAAGGTCACTCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTC
 GGCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCC
 CCTGGTCAAGACCATCGTGGAGTTCCACATGACGACTGAGGCCAAGCCACCATCCGCATGGA
 CACCAGTGCAAGTGGCCCCACCCGCCTGGTCCTCAGTGACTGTGCCACCAGCCATGGGAGCCT
 GCGCATCCAAGTGTATAAGCTCTCCTGGTGAACGCCCTAGCTAACAGGTCACTGAA
 CCTCCTAGTGCCATCCCTGCCAATCTAGTAAAAACCAGCTGTGCCCAGTCAGGCTTC
 CTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTAAGGTGCCATTCCCTCAGCATTGA
 CCGTCTGGAGTTGACCTTCTGTATCCTGCCATCAAGGGTACACCATTAGCTTACCTGGG
 GGCAAGTTGGACTCACAGGAAAGGTGACCAAGTGGTCAATAACTCTGCAGCTCCCT
 GACAATGCCACCCTGGACAACATCCCGTTAGCCTCATCGTGAGTCAGGACGTGGTAAAGC
 TGCAGTGGCTGCTGTGCTCTCCAGAAGAATTCACTGGCCTGGACTCTGTGCTTCTGA
 GAGTGCCATCGGCTGAAGTCAAGCATCGGCTGATCAATGAAAAGGCTGCAGATAAGCTGGG
 ATCTACCCAGATCGTGAAGATCTAACACTCAGGACACTCCGAGTTTATAGACCAAGGCCA
 TGCCAAGGTGCCCAACTGATCGTGGACTGGAAAGTGGTCCCTCCAGTGAAGCCCTCCGCC
 GTTCACCCTGGCATCGAAGCCAGCTCGGAAGCTCAGTTACACCAAGGTGACCAACTTAT
 ACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTGATGAACCTGGGATTGGCTGGT
 CCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCACTCCATCCTGCTGCCAACCA
 GAATGGCAAATTAAGATCTGGGGTCCCAGTGTCAATTGGTGAAGGCCTGGGATTGAGGCAGC
 TGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCTGTGGAAACCCAGCTC
 TCCTGTCTCCCAG**TGA**AGACTTGGATGGCAGCCATCAGGAAAGGCTGGTCCCAGCTGGAGT
 ATGGGTGTGAGCTATAGACCATCCCTCTGCAATCAATAACACTTGCCTGTGAAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQPLLSAMR
EKPGAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVK
TIVEFHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLV
PSLPNLVKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKL
LDSQGKVTKWFNNSAASLTMPTLDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLLDSVLPEAH
RLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEALRPLFTL
GIEASSEAQFYTKGDQLILNLNNISSLRIQLMNSIGWFQPDVLKNIITEIIHSILLPNQNGK
LRSGVPVSLVKALGFEAAESSLTKDALVLTPASLWKPPSPVVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGA
 GCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGC
 CTCTCTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTCTGGGGCTTTGGGCACACTGGT
 TGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTATGTCGGTGCCAGCATTGTGACAGCAGT
 TGGCTTCTCCAAGGGCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGA
 CATCTATAGCACCCCTCTGGGCCTGCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGAC
 ATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGCATGAGATGCACAGTCTT
 CTGCCAGGAATCCCGAGCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTCATCCTTGG
 AGGCCTCCTGGGATTCTTCCTGTTGCCTGGAATCTTCATGGATCCTACGGACTTCTACTC
 ACCACTGGTGCCTGACAGCATGAAATTGAGATTGGAGAGGGCTTTACTTGGCATTATTC
 TTCCCTGTTCTCCCTGATAGCTGGAATCATCCTCTGCTTTCTGCTCATCCCAGAGAAATCG
 CTCCAACTAACGATGCCAACCAAGCCAAACCTCTGCCACAAGGAGCTCTCCAAGGCCTGG
 TCAACCTCCAAAGTCAAGAGTGAGTTCAATTCTACAGCCTGACAGGGTATGTGTGAAGAAC
 CAGGGGCCAGAGCTGGGGGTGGCTGGCTGTGAAAAACAGTGGACAGCACCCGAGGGCCA
 CAGGTGAGGGACACTACCCTGGATCGTGTAGAAGGTGCTGCTGAGGATAGACTGACTTGG
 CCATTGGATTGAGCAAAGGCAGAAATGGGGCTAGTGTAAACAGCATGCAGGTTGAATTGCCAA
 GGATGCTGCCATGCCAGCCTTCTGTTCTCACCTGCTGCTCCCCTGCCCTAAGTCCCC
 AACCTCAACTTGAAACCCATTCCCTTAAGCCAGGACTCAGAGGATCCCTTGCCCTGTT
 TTACCTGGACTCCATCCCCAAACCCACTAACATCACATCCCAGTGAACCCCTCTGTGATCAA
 AGACCCTCTCTGGCTGAGGTTGGCTTAGCTCATTGCTGGGATGGGAAGGAGAACAGT
 GGCTTTGTGGCATTGCTCTAACCTACTTCTCAAGCTTCCCTCAAAGAAACTGATTGGCCC
 TGGAACCTCCATCCACTCTGTATGACTCCACAGTGTCCAGACTAATTGTGCATGAACTG
 AAATAAAACCATCCTACGGTATCCAGGAAACAGAAAGCAGGATGCAGGATGGGAGGACAGGAA
 GGCAGCCTGGGACATTAAAAAAATA

FIGURE 80

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQ
CDIYSTLLGLPADIQAAQAMMVTSSAISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFI
LGGLLGFIIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIIILCFSCSSQR
NRSNYYDAYQAQPLATRSSPRPGQPPKVKEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTCTGCTGGACCTTCCTCGTCTCTCCATCTCTCCCTCCTT
 TCCCCGCGTTCTCTTCCACCTTCTCTTCCCACCTTAGACCTCCCTGCCCTCCT
 TCCTGCCAACCGCTGCTCCTGGCCCTCTCCGACCCGCTCTAGCAGCAGACCTCCTGGGGT
 CTGTGGGTTGATCTGTGGCCCTGTGCCTCCGTGTCCCTTCGTCTCCCTCCCAGACTCC
 GCTCCGGACCAGCGGCCTGACCCTGGGAAAGGATGGTTCCCGAGGTGAGGGTCCCTCCTC
 CTTGCTGGACTCGCGCTGCTCTGGTCCCCCTGGACTCCCACGCTCGAGCCGCCAGACAT
 GTTCTGCCTTTCCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCC
 ACAAGGCCTGATGTACTGCCTGCCTGTACCTGCTCAGAGGGGCCATGTGAGTTGTTACCG
 CCTCCACTGTCCGCCTGTCCACTGCCCTGCCTGTGACGGAGCCACAGCAATGCTGTCCCAA
 GTGTGTGGAACCTCACACTCCCTCTGGACTCCGGGCCCCACCAAAGTCCTGCCAGCACACGG
 GACCATGTACCAACACGGAGAGATCTCAGTGCCCATGAGCTGTTCCCCTCCGCCCTGCCCAA
 CCAGTGTGTCCTCTGCAGCTGCACAGAGGGCCAGATCTACTGCCCTCACAAACCTGCCCGA
 ACCAGGCTGCCAGCACCCCTCCACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGC
 AAGTGAGCAATCGGATGAAGAGGACAGTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGA
 TCCATGTTCCAGTGTGCTGGAGAAAGAGAGAGGCCGGCACCCAGCCCCACTGGCCTCAG
 CGCCCCCTTGAGCTTCATCCCTGCCACTTCAGACCCAAGGGAGCAGGCAGCACACTGTCAA
 GATCGTCCTGAAGGAGAACATAAGAAAGCCTGTGTGCATGGCGGGAAAGACGTACTCCCACGG
 GGAGGTGTGGCACCCGGCTTCCGTGCCTTCGGCCCTTGCCCTGCATCCTATGCACCTGTGA
 GGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCCCACCGAGTACCCCTGCCGTACCCCGA
 GAAAGTGGCTGGGAAGTGCTGCAAGATTGCCAGAGGACAAAGCAGACCCCTGCCACAGTGA
 GATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGTCCTCGTCCACACATCGGTATCCCC
 AAGCCCAGACAACCTGCGTCGCTTGCCCTGGAACACAGAGGCCTCGGACTTGGTGGAGATCTA
 CCTCTGGAAGCTGGTAAAAGATGAGGAAACTGAGGCTCAGAGAGGTGAAGTACCTGGCCCAAG
 GCCACACAGCCAGAACATCTCCACTTGACTCAGATCAAGAAAGTCAGGAAGCAAGACTTCCAGA
 AAGAGGCACAGCACTCCGACTGCTCGCTGGCCCCACGAAGGTCACTGGAACGTCTTCCTAG
 CCCAGACCCCTGGAGCTGAAGGTACGGCCAGTCCAGACAAAGTGACCAAGACATAACAAAGAC
CTAACCAGTTGCAGATATGAGCTGTATAATTGTTATTATATTAAATAAAATAAGAAGTTGC
 ATTACCCCTCAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTC
SEGAHVSCYRLHCPPVHCPQPVTETPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSA
HELFPSRLPNQCVLCSCTEGQIYCGLTTCPPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQ
SLHGVRHPQDPCSSDAGRKRGPGTAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEHKKAC
VHGGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCCKICP
EDKADPGHSEISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETE
AQRGEVPGPRPHSQNLPLDSDQESQEARNPERGTALPTARWPPRRSLERLPSPDPGAEGHGQS
RQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCCGCTCACGCAGA
 GCCTCTCCGTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTCTCTAATCCATCC
 GTCACCTCTCCTGTCATCCGTTCCATGCCGTGAGGTCCATTACAGAACACATCC**ATG**GCTC
 TCATGCTCAGTTGGTTCTGAGTCTCCTCAAGCTGGATCAGGGCAGTGGCAGGTGTTGGC
 CAGACAAGCCTGTCCAGGCCTGGTGGGGAGGACGCAGCATTCTCCTGTTCTGTCTCCTA
 AGACCAATGCAGAGGCCATGGAAGTGCAGGTTCTCAGGGCCAGTTCTCTAGCGTGGTCCACC
 TCTACAGGGACGGGAAGGACCAGCCATTATGCAGATGCCACAGTATCAAGGCAGGACAAAAC
 TGGTGAAGGATTCTATTGCGGAGGGCGCATCTCTGAGGCTGGAAAACATTACTGTGTTGG
 ATGCTGGCCTCTATGGGTGCAGGATTAGTCCCAGTCTTACTACCAGAACGCCATCTGGGAGC
 TACAGGTGTCAGCACTGGGCTCAGTCCTCTCATTCATCACGGATATGTTGATAGAGACA
 TCCAGCTACTCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTCCAC
 AAGGACAGGATTGTCCACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTGATGTGG
 AGATCTCTCTGACCGTCCAAGAGAACGCCGGAGCATATCCTGTTCCATGCCATGCTCATC
 TGAGCCGAGAGGTGGAATCCAGGGTACAGATAGGAGATACCTTTGAGCCTATATCGTGGC
 ACCTGGCTACCAAAGTACTGGGAAATACTCTGCTGTGGCCTATTTTGGCATTGTTGGACTGA
 AGATTTCTTCTCCAAATTCCAGTGGAAAATCCAGGCGGAACGGACTGGAGAAGAAAGCACG
 GACAGGCAGAATTGAGAGACGCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGG
 CTCACCGAAGCTCTGCGTTCTGATCTGAAAACACTGTAACCCATAGAAAAGCTCCCCAGGAGG
 TGCCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAG
 GGAAACATTACTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCGTGGAGTGTGCCGG
 ATGATGTGGACAGGAGGAAGGAGTACGTGACTTTGTCTCCGATCATGGTACTGGTCTCCTCA
 GACTGAATGGAGAACATTGTATTCACATTAAATCCCCGTTTATCAGCGTCTCCCCAGGA
 CCCCACCTACAAAAATAGGGTCTCCTGGACTATGAGTGTGGACCATCTCCTCTTCAACA
 TAAATGACCAGTCCCTTATTACCCCTGACATGTCGGTTGAAGGCTTATTGAGGCCCTACA
 TTGAGTATCCGTCTATAATGAGCAAAATGGAACCTCCATAGTCATCTGCCAGTCACCCAGG
 AATCAGAGAAAGAGGCCTCTGGCAAAGGGCTCTGCAATCCCAGAGACAAGCAACAGTGAGT
 CCTCCTCACAGGCAACCACGCCCTCCTCCCCAGGGTGAAATG**TAG**GATGAATCACATCCCA
 CATTCTCTTAGGGATATTAAGGTCTCTCTCCCAGATCCAAAGTCCCGCAGCAGCCGGCAA
 GGTGGCTTCCAGATGAAGGGGACTGGCCTGTCCACATGGAGTCAGGTGTCATGGCTGCCCT
 GAGCTGGGAGGGAAGAAGGCTGACATTACATTAGTTGCTCTCACTCCATCTGGCTAAGTGA
 TCTTGAAATACCACCTCTCAGGTGAAGAACCGTCAGGAATTCCATCTCACAGGCTGTGGTGT
 AGATTAAGTAGACAAGGAATGTGAATAATGCTTAGATCTTATTGATGACAGAGTGTATCCTAA
 TGGTTGTTCATTATATTACACTTCACTAA

FIGURE 84

MALMLSLVSLKLGSQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSV
VHLYRDGKDQPFMOMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYYQKAI
WELQVSALGSVPPLISITGYVDRDIQLLCQSSGWPRPTAKWKGPQGQDLSTDRTNRDMHGLF
DVEISLTQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIV
GLKIFFSKFQWKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAP
QEVPHSEKRFRTRKSVVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYW
VLRLNGEHLYFTLNPRFISVFPPRKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLR
PYIEYPSYNEQNGTPIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCCAGAC**ATG**CTGCTGCTGCTGCC
 CTGCTCTGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCC
 GTGACGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCTCCTACCCCTCGCATGGC
 TGGATTACCCTGGCCCAGTAGTCATGGCTACTGGTTCCGGGAAGGGCCAATACAGACCAG
 GATGCTCCAGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGAGGAGACTCGGGACCGATT
 CACCTCCTGGGGACCCACATACCAAGAATTGCACCCCTGAGCATTAGAGATGCCAGAAGAAGT
 GATGCGGGAGATACTTCTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCAC
 CGGCTCTGTGAATGTGACAGCCTGACCCACAGGCCAACATCCTCATCCCAGGCACCC
 GAGTCGGCTGCCCTAGAATCTGACCTGCTCTGTGCCCTGGCCTGTGAGCAGGGACACCC
 CCTATGATCTCCTGGATAGGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCG
 GTGCTCACCCATCCCACAGCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTC
 CCTGGGCCAGCGTGACCACGAACAAGACCGTCCATCTAACGTGTCCTACCCGCTCAGAAC
 TTGACCATGACTGTCTCCAAGGAGACGGCACAGTATCCACAGTCTGGAAATGGCTCATCT
 CTGTCACTCCCAGAGGCCAGTCTGCGCCTGGTCTGTGAGTTGATGCAGTTGACAGCAAT
 CCCCCGCCAGGCTGAGCCTGAGCTGGAGAGGCCCTGACCTGTGCCCTCACAGCCCTCAAAC
 CCGGGGGTGCTGGAGCTGCCTTGGGTGCACCTGAGGGATGCAGCTGAATTCACCTGCAGAGCT
 CAGAACCCCTCTGGCTCTCAGCAGGTCTACCTGAACGTCTCCCTGCAGAGCAAAGCCACATCA
 GGAGTGACTCAGGGGTGGTCGGGGAGCTGGAGCCACAGCCCTGGTCTTCCTGTCCTCTGC
 GTCATCTCGTTGAGGTCTGCAGGAAGAAATCGGCAAGGCCAGCAGCGGGCTGGGA
 GATAACGGGCATAGAGGATGCAAACGCTGTCAGGGTTCAAGCTCTCAGGGGCCCTGACTGAA
 CCTTGGGCAGAACAGACTCCAGACCCAGACAGCCTCCCCAGCTTCTGCCGCTCCTCAGTGGGG
 GAAGGAGAGCTCCAGTATGCATCCCTCAGCTCCAGATGGTGAAGCCTGGACTCGCGGGGA
 CAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAGA**TGA**GAAACTGCAGAGACT
 CACCCCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCAGAGGCTGATTCTGTAG
 AATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATGTGCAGAGTGAAAAGC
 ACACAGGCTTAGAGTCAGGAAACTCTCAAACCTGAATCCACACTGTGCCCTCCCTTTATT
 TTTAACTAAAAGACAGACAAATTCTA

FIGURE 86

MLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPSCFSYPSHGWIYPGPVVHGYWFR
EGANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLISRSDAGRFFRMEKGSI
KWNYKHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSPWACEQGTPPMISWIGTSVSPLD
PSTTRSSVLTLIPQPQDHGTSLTCQVTFPGASVTTNKTVHLNVSYPPQNLTMVFQGDGTVST
VLGNNGSSLSLPEGQLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDA
AEFTCRAQNPLGSQQVYLNVLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSA
RPAAGVGDTGIEDANAVRGSAQGPLTEPWAEDSPPDQPPPASARSSVGEHELQYASLSFQMV
KPWDSRGQEATDTEYSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTGT
ACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACA**ATG**AACCAACTCAGCTTCCTGCTG
TTTCTCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGG
ACCTGTTCTCGTCTCCATCTCTGCCAGAACGCTGCAAGGAAATCAAAGACGAATGTCCTAGT
GCATTGATGGCCTGTATTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGAC
ATGACCTCTGGGGTGGCGGCTGGACCCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGG
AAGTGCACGGTGGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGG
GACGGCAACTGGGCCAACTACAACACCTTGGATCTGCAGAGGCCACGAGCGATGACTAC
AAGAACCCCTGGCTACTACGACATCCAGGCCAAGGACCTGGCATCTGGCACGTGCCAATAAG
TCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTC
CAGACACTGGGACATAATCTGTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGA
AAGTGTGGACTGACAACGGCCCGGTGATCCCTGTGGTCTATGATTTGGCGACGCCAGAAA
ACAGCATCTTATTACTCACCCATGGCCAGCGGGATTCACTGCAGGGATTGTTCAAGTTCAG
GTATTTAATAACGAGAGAGCAGCCAACGCCCTGTGTGCTGGAATGAGGGTCACCGGATGTAAC
ACTGAGCATCACTGCATTGGTGGAGGAGGATACTTCCAGAGGCCAGTCCCCAGCAGTGTGGA
GATTTTCTGGTTTGATTGGAGTGGATATGGAACTCATGTTGGTTACAGCAGGCCGTGAG
ATAACTGAGGCAGCTGTGCTTCTATTCTATCGT**TGA**GAGTTGTGGAGGGAACCCAGACCT
CTCCTCCCAACCATGAGATCCCAAGGATGGAGAACAACTTACCCAGTAGCTAGAATGTTAATG
GCAGAAGAGAAAACAATAAATCATATTGACTCAAGAAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSPSLPRSCKEIKDECPSAFDGLYFLRTENG
VIYQTFCDMTSGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSA
EAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQK
YPVKYGECKCWTDNGPVI PVVYDFGDAQKTASYYSPYGQREFTAGFVQFRVFNNERAANALCA
GMRVTGCNTEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTTGCAGGGAGACTTCAGGAGTCGCTGTCTGAACCCAGCCTCAGAG
ACCGCCGCCCTTGTCCCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCT
GACGCTCCTGGCGCATCTGGTGGTCGTACACCTTATTCTGGTCCCAGGACAGAACATACA
GGCCTGCCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCACTGGTGGC
CGCGCTCTCTGTCACCCTGGGCCTTTGCAGTGGAGCTGCCGGTTCTCAGGAGTCTC
CATGTTAACACAGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCT
GTCCTTCTTCATATTGAGCGTTGGAGTGCACGTATTGGTACATTGTCTGCAG
TGCCCTTCCAGCTGTCACTGAAATGGCTTATTGTCACCGTCTTGGCTGAAAAAGAAACC
CTTCTGATTACCTTCATGACGGAACCTAACGGACGAAGCCTACAGGGCAAGGGCCGCTCGT
ATTCCCTGGAAGAAGGAAGGCATAGGCTTCGGTTTCCCCTCGAAACTGCTCTGCTGGAGGA
TATGTGTTGGAATAATTACGTCTGAGTCTGGGATTATCCGCATTGTATTAGTGCCTTGAA
TAAAATATGTTTAGTAACATTAAGACTTACAGTTAGGGACAATTAAAAAAA
AAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGL
FAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEM
ALFVTVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAG**ATG**CAACTG
 ACTCGCTGCTGCTCGTGTTCCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGCCAGGAT
 GATGGCCTCCCGGCTCAGAGGACCCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGG
 GTGCCTCGGAAGCGGGGCCACATCTCACCTAACCTAAGTCCCCTCATGCCATTCCACTCTCCTA
 GGGCTGCTGGCCCCGCCTGGGGAGGGCTTGGGCATTCTGGCAGCCCCCAACCGCCCGAAC
 CACAGCCCCCACCCTCAGCCAAGGTGAAGAAAATTTGGCTGGGCAGCTACTCCAAC
 ATCAAGACGGTGGCCCTGAACCTGCTCGTCACAGGGAAAGATTGTGGACCATGGCAATGGACC
 TTCAGCGTCCACTTCAAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGC
 CCCAGTAAAGCTGTAGAGTTCCACCAGGAACACAGCAGATCTTCATCGAACGCCAAGGCCTCCAAA
 ATCTTCAAATGCCGGATGGAGTGGAGAAGGTAGAACGGGCCGGACCTCGCTTGCACC
 CACGACCCAGCCAAGATCTGCTCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCC
 CAGCCCTCAAAGTCGTCTGTCTACATGCCCTCTACAGCACGGACTATCGGCTGGTCCAG
 AAGGTGTGCCAGATTACAACCTACCATAGTATACTACCCCTACTACCCATCTGGG**TGA**CCCGGG
 GCAGGCCACAGAGGCCAGGGCTGGAGGACAGGCCTGCCATGCAGGAGACCATCTGG
 ACACCGGGCAGGGAAAGGGTTGGGCCTCAGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAA
 GTGGGCCAGGGCCAAGTCTCAAGTGGCAGAGAAAGGGTCCAAGTGCTGGTCCAACCTGAA
 GCTGTGGAGTGAATCACAGGAGCACTGGAGGAGGTGGCTCTGTGCAGCCTCACA
 GGGCTTGCCACGGAGCCACAGAGAGATGCTGGTCCCCGAGGCCTGTGGCAGGCCGATCAG
 TGTGGCCCCAGATCAAGTCATGGAGGAAGGCTAACGCCCTGGTCTGCCATCCTGAGGAAAG
 ATAGCAACAGGGAGGGGAGATTCTACAGTGTGGACAGCCTGTCAACTTAGGATGGATGGCT
 GAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGGCCAGAGGAGCTCTCCAGCCCTG
 CCTAGTGGCGCCCTGAGCCCCTGTGCTGTGCTGAGCATGGCATGAGGCTGAAGTGCAACC
 CTGGGGTCTTGATGTCTTGACAGATTGACCATCTGTCTCCAGCCAGGCCACCCCTTCCAAA
 ATTCCCTCTGCCAGTACTCCCCCTGTACCAACCCATTGCTGATGGCACACCCATCCTTAAG
 CTAAGACAGGACGATTGTGGCCTCCCACACTAACGCCACAGCCATCCGCGTGTGTGTC
 CCTCTCCACCCCAACCCCTGCTGGCTCCTCTGGAGCATCCATGTCCGGAGAGGGGTC
 CAACAGTCAGCCTCACCTGTCAGACCAGGGTTCTCCGGATCTGGATGGCAGGCCCTCTCAG
 CAGCGGGCACGGTGGGGGGCCGGCGAGAGCATGTGCTGGATCTGTTCTGTGTCT
 GTCTGTGGGTGGGGGGAGGGAGGGAAAGTCTGTGAAACCGCTGATTGCTGACTTTGTGTGA
 AGAATCGTGTCTGGAGCAGGAAATAAGCTGCCCGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMA
TLLGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHG
NGTFSVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIIEAKASKIFNCRMEWEKVERGRRTS
LCTHDPAKICSRDHAQSSATWSCSQPKVVVCVYIAFYSTDYRLVQKCPDNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTCGGCTGCCTTCATTGCCTCGGGCCTGCGCTCG
CCCTTATGTCTCACCATGCCATCGAGCCGTTGCGTATCATCTCCTCATGCCGGAGCTT
TCTTCTGGTTGGTGTCTACTGATTCGTCCCTGTTGGTCATGGCAAGAGTCATTATTG
ACAACAAAGATGGACCAACACAGAAATATCTGCTGATCTTGGAGCGTTGTCTGTCTATA
TCCAAGAAATGTTCCGATTGCATATTATAAACTCTTAAAAAAAGCCAGTGAAGGTTGAAGA
GTATAAACCCAGGTGAGACAGCACCCCTATGCGACTGCTGGCCTATGTTCTGGCTGGGCT
TTGGAATCATGAGTGGAGTATTCCTTGTAATACCCTATCTGACTCCTGGGCCAGGCA
CAGTGGCATTCATGGAGATTCTCCTCAATTCTCCTTATTCACTTGCATGACGCTGGTCA
TTATCTGCTGCATGTATTCTGGGCATTGTATTTTGATGGCTGTGAGAAGAAAAAGTGGG
GCATCCTCCTTATCGTTCTCCTGACCCACCTGCTGGTGTAGCCCAGACCTTCATAAGTTCTT
ATTATGGAATAAACCTGGCGTCAGCATTATAATCCTGGTGCTCATGGCACCTGGCATTCT
TAGCTGCGGGAGGCAGCTGCCGAAGCCTGAAACTCTGCCTGCTGCCAAGACAAGAACTTTC
TTCTTACAACCAGCGCTCCAGATAACCTCAGGGAACCAGCACTTCCAAACCGCAGACTACA
TCTTAGAGGAAGCACAACTGTGCCTTTCTGAAAATCCCTTTCTGGTGGATTGAGAAA
GAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSSLISSLVWFMARVIIDNK
DGPTQKYLLIFGAFVSFYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGI
MSGVFSFVNTLSDSLGP GTVG I HGDSPQFFLYSAFMTLVIIILLHVFWGIVFFDGCEKKKGIL
LIVLLTHLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKLCLLCQDKNFLLY
NQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCG
TGGACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTGTCTTAGGAT
CAAACTCGGTCATTACCACAGCTCAAACCTGCTTGGACTCCCTCCCACAAAATGGCTCCGG
ATCAGGGAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTCCTTAAAGTCTGATAC
CATTAACACAGATGCTCACACTGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGA
CACCTGGTACCCAGACCCACCCATTGACCTGGGAGGGTTGAATGTACAACAGCAACTGCACC
CACATGTGTTACCAATTTGTCACACAACACTGGAGGCCAGGGCACTATCCTAAGCTCAGAGG
AATTGCCACAAATCTCACGAGCCTCATCCATTGTTCCGGAGGCATCCTGCCCA
CCAGTCAGGCAGGGCTAATCCAGATGTCCAGGATGGAAGCCTCCAGCAGGAGGAGCAGGTG
TAAATCCTGCCACCCAGGAACCCCAGCAGGCCCTCCCAACTCCAGTGGCACAGATGACG
ACTTTGCAGTGACCACCCCTGCAGGCATCCAAGGAGCACACATGCCATCGAGGAAGCCACCA
CAGAATCAGCAAATGGAATTCAGTAAGCTGTTCAAATTGAAACTAAGCTGCCTCGAATT
TGGTGATACTGTGAATCTTATCATTGATTATATTGGAATAGATTGAGACACATTGGATA
GTCTTAGAAGAAATTAAATTCTAATTACCTGAAAATATTCTGAAATTGAGAAAATATGTT
CTATGTAGAGAATCCAACTTTAAAAACAATAATTCAATGGATAATCTGTCTTGAAATAT
AACATTATGCTGCCTGGATGATGCATATTAAACATATTGGAAAATGGAAAAA
AAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSSLIPLTQM
LTLGPDLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIIEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCTGCCCTGCCCTCCCCACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCT
 CTCTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCCTGGCGCTGGAGGGCCTGTCCCTGAC
CATGGTCCCTGCCTGGCTGTGGCTGTTGTCTCCGTCCCCCAGGCTCTCCCCAAGGCCA
 GCCTGCAGAGCTGTCTGTGGAAGTCCAGAAAATCTGGTGGAAATTCCCTTATACCTGAC
 CAAGTTGCCGCTGCCCGTGAGGGGGCTGAAGGCCAGATCGTGCAGGGACTCAGGCAA
 GGCAACTGAGGGCCCATTGCTATGGATCCAGATTCTGGCTCCTGCTGGTGCACCAGGCCCT
 GGACCGAGAGGAGCAGGCAGAGTACCTACAGGTACCCCTGGAGATGCAGGATGGACATGT
 CTTGTGGGTCCACAGCCTGTGCTGCACGTGAAGGATGAGAATGACCAGGTGCCCTT
 CTCTCAAGCCATCTACAGAGCTCGGCTGAGCCGGGTACCAAGGCCTGGCATCCCCCTCCT
 CCTTGAGGCTTCAGACCGGGATGAGCCAGGCACAGCCAACCTGGATCTCGATTCCACATCCT
 GAGCCAGGCTCCAGCCCAGCCTCCCCAGACATGTTCCAGCTGGAGCCTGGCTGGGGCTCT
 GGCCCTCAGCCCCAAGGGGAGCACCAGCCTGACCACGCCCTGGAGAGGACCTACCAGCTGTT
 GGTACAGGTCAAGGACATGGGTGACCAGGCCTCAGGCCACAGGCCACTGCCACCAGTGGAAAGT
 CTCCATCATAGAGAGCACCTGGGTGTCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGT
 CCTATAACCGCACACATGGCCAGGTACACTGGAGTGGGGTAGTGTGCACTATCACCTGGA
 GAGCCATCCCCGGGACCCTTGAAAGTGAATGCAGAGGGAAACCTCTACGTGACCAGAGAGCT
 GGACAGAGAACGCCAGGCTGAGTACCTGCTCAGGTGCAGGCTCAGAATTCCATGGCGAGGA
 CTATGCGGCCCTCTGGAGCTGCACGTGCTGGTAGGGATGAGAATGACAACGTGCCTATCTG
 CCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCTCAGTCCACCAGGTACTGAAGTGACTAG
 ACTGTCAAGCAGAGGATGCAGATGCCCGGCTCCCCAATTCCACGTTGTATCAGCTCCT
 GAGCCCTGAGCCTGAGGATGGGTAGAGGGGAGAGCCTCCAGGTGGACCCCACCTCAGGCAG
 TGTGACGCTGGGGTGCTCCACTCCGAGCAGGCCAGAACATCCTGCTCTGGTGCCTGGCCAT
 GGACCTGGCAGGCGCAGAGGGTGGCTCAGCAGTCAGTGAAGTCGAAGTCGCAGTCACAGA
 TATCAATGATCACGCCCTGAGTTCATCACTCCCAGATTGGCCTATAAGCCTCCCTGAGGA
 TGTGGAGCCCCGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGACCTCGAGCCCGCCT
 CGCCTCATGGATTTCGCCATTGAGAGGGAGACACAGAACGGACTTTGGCCTGGATTGGGA
 GCCAGACTCTGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAGGCAGCTCCAAG
 TCATGAGGTGGTGGTGGTAGCAGAGTGTGGGAAGCTGGTGGGGCCAGGCCAGGCCCTGG
 AGCCACCGCCACGGTACTGTGCTAGTGGAGAGAGTGTGGCCACCCCCAAGTGGACCAGGA
 GAGCTACGAGGCCAGTGTCCCCATCAGTGCCTCAGGCCCTGACTCTGCCCTGTGCCCTCCA
 CTCCGACCCATCAGCCGAACCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCTCTG
 CATTGAGAAATTCTCCGGGAGGTGCACACGCCAGTCCCTGCAGGGCCTGGCTGGGG
 CACCTACACGGTGTGTTGGAGGCCAGGATAACGCCCTGACTCTGCCCTGTGCCCTCCA
 ATACCTCTGCACACCCGCCAAGACCATGGCTGATCGTGAGTGGACCCAGCAAGGACCCGA
 TCTGGCCAGTGGCACGGTCCCTACAGCTTACCCCTGGTCCAAACCCACGGTGCAACGGGA
 TTGGCGCTCCAGACTCTCAATGGTCCCATGCCCTACCTCACCTGGTCCAAACCCACGGTG
 GCCACGTGAACACATAATCCCCTGGTAGCCACAATGCCAGATGTGGCAGCTCCTGGT
 TCGAGTGATCGTGTGCGTCAACGTGGAGGGCAGTGCATGCGCAAGGTGGCCGCATGAA
 GGGCATGCCACGAAGCTGCGCAGTGGCATCCTGTAGGCACCCCTGGTAGCAATAGGAAT
 CTTCCCTCATCCTCATTTCACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGC
 AGACAGCGTGCCCTGAAGGCCAGTGTCTGAAGGCGACTGTCT**TGA**ATGGCCCAGGCAGCTCTAGCTGGGAGCTTGGC
 CTCTGGCTCCATCTGAGTCCCTGGGAGAGAGGCCAGCACCAAGATCCAGCAGGGACAGGA
 CAGAGTAGAAGCCCTCCATCTGCCCTGGGTGGAGGCACCATCACCATCACCAAGGCATGTCT
 GCAGAGCCTGGACACCAACTTATGGACTGCCCATGGAGTGCTCAAATGTCAGGGTGTGG
 CCCAATAATAAGCCCCAGAGAACTGGCTGGCCCTATGGAAAAA
 AAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPYLTKLPLPREGAEGQIVLSGDSGK
ATEGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLHVKDENDQVPHF
SQAIYRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSPDMFQLEPRLGAL
ALSPKGSTS LDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLEPIHLAENLKV
LYPHHMAQVHWSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGED
YAAPPLEHVLVMDENDNPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQOLL
SPEPEDGVGRAFQVDPTSGSVTLGVPLRAGQNILLVLAMDLAGAEGGSSTCEVEVAVTD
INDHAPEFITSQIGPISLPEDVEPGTLVAMLTAIDADLEPAFRIMDFAIERGDTEGTFGLDWE
PDSGHVRLRLCKNLSYEAPSHEVVVVVQSVAKLGVPGPGPGATATVTVLVERVMPPPQLDQE
SYEASVPISAPAGSFLLTIQPSDPISRTLRFSLVNDSEGWL CIEKFSGEVHTAQSLQGAQPGD
TYTVLVEAQDTALT LAPVPSQYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRD
WRLQTLNGSHAYLTLALHWVEPREHIIPVVVSHNAQM WQLLVRVIVCRCNVEGQCMRKVGRMK
GMPTKLSAVGILVGTLVAIGIFLILIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCACGCC
 TGAGTCCAAGATTCTCCCAGGAACACAAACGTAGGAGACCCACGCTCTGGAAGCACCAGCC
 TTTATCTCTCACCTCAAGTCCCCTTCTCAAGAACATCCTCTGTTCTTGCCTCTAAAGTCT
 TGGTACATCTAGGACCCAGGCATCTGCTTCCAGCCACAAAGAGACAGATGAAAGATGCAGAA
 AGGAAATGTTCTCCTATGTTGGTCTACTATTGCATTAGAAGCTGCAACAAATTCCAATGA
 GACTAGCACCTCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACGCCACCAA
 CTCTGGGTCCAGTGTGACCTCCAGTGGGTCAAGCACAGCCACCATCTCAGGGTCCAGCGTGAC
 CTCCAATGGGTCAAGCATAGTCACCAACTCTGAGTTCCATACAACCTCCAGTGGGATCAGCAC
 AGCCACCAACTCTGAGTTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTC
 CAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGG
 GCCAGCACAGTCACCAACTCTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAA
 CTCTGAGTCCAGCACAGTCAGTAGGGCCAGCAGTCCACAGCACACCCACTCTGAGTCTAGCACACT
 CTCCAGTGGGCCAGCACAGCCACCAACTCTGACTCCAGCACAAACCTCCAGTGGGCTAGCAC
 AGCCACCAACTCTGAGTCCAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTC
 CAGCACAGTGTCCAGTAGGGCCAGCAGTCCACAGCACACCCACTCTGAGTCCAGCACACCCTCCAGTGG
 GCCAGCACAGCCACCAACTCTGAGTCCAGAACGACCTCAATGGGCTGGCACAGCCACCAA
 CTCTGAGTCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGACTCCAGCACAGT
 GTCCAGTGGGCCAGCAGTCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGCCAGCAC
 AGCCACCAACTCTGAGTCCAGCACACCCACTCTGAGTCCAGCACAGCCACCAACTCTGACTC
 CAGCACAAACCTCCAGTGGGCCGGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTGG
 GATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCACTCCAGTGGGCCAACACAGCCACCAA
 CTCTGAGTCCAGTACGACCTCCAGTGGGCCAACACAGCCACCAACTCTGAGTCCAGCACAGT
 GTCCAGTGGGCCAGCAGTCCACCAACTCTGAGTCCAGCACACCCACTCCAGTGGGCCAGCAC
 AGCCACCAACTCTGAGTCCAGCACAAACCTCCAGTGGGCTAGCACAGCCACCAACTCTGACTC
 CAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAGCACAGTGTCCAGTGG
 GATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCACTCCAGTGGGCCAACACAGCCACCAA
 CTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGAATGCACACAAC
 TTCCCATAGTCATCTACTGCAGTGAGTGAGGCAAAGCCTGGTGGTCCCTGGTCCGTGGGA
 AATCTTCCTCATCACCCTGGTCTCGGTTGTGGCGGCCGTGGGCTCTTGCTGGCTCTTCT
 CTGTGTGAGAACACAGCCTGTCCTGAGAACACACCTTAACACAGCTGTCTACCACCCCTCATGG
 CCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAGGTG
 GAGTCCTAACTGGTTCTGGAGGGAGACCAAGTATCATCGATGCCAGGGCTGGGAGATGAGCAGGAA
 CAGCGGGCCTGAGCAGCCCCGGAAAGCAAGTGCCGCATTCTCAGGAAGGAAGAGACCTGGGC
 ACCCAAGACCTGGTTCTTCATTCATCCCAGGAGACCCCTCCAGCTTGTGAGATCCT
 GAAAATCTTGAAGAACAGGTATTCTCACCTTCTGCCTTACCAAGACACTGGAAAGAGAACATAC
 TATATTGCTCATTAGCTAACGAAATAAACATCTCATCTAACACACACGACAAAGAGAACAGT
 GTGCTTGCCTGGGGTGGGTATCTAGCTCTGAGATGAACACTAGTTATAGGAGAAACCTCCAT
 GCTGGACTCCATCTGGCATTCAAATCTCCACAGTAAAATCCAAAGACCTCAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMKGKVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATIS
GSSVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESS
TPSSGASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTS
SGASTATNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGA
GTATNSESSTTSSGASTATNDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTA
TNSDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNS
ESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTATNDSSTSSEASTATNSESS
TVSSGISTVTNSESSTTSSGANTATNSGSSVTSAGSGTAALTGMHTSHSASTAVSEAKPGGS
LVPWEIFLITLVSVVAAVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAP
HRPRWSPNWFWRPVSSIAMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCGGGACGCCTCCGCGTTACGGGATGAATTAAACGGCGGGTCCGCACGGAGGTTGTGACCCC
 TACGGAGCCCCAGCTGCCACGCACCCCACTCGGCGCGCGCGTGCCTGCTGTACACA
 GGTGGGAGGCTGGAACTATCAGGCTGAAAAACAGAGTGGGTACTCTCTGGGAAGCTGGCA
 ACAAAATGGATGATGTGATAT**ATG**CATTCCAGGGAAAGGGAAATTGTGGTGCTTCTGAACCCAT
 GGTCAATTAACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCT
 TTGGAATCATGGTGTATGGAAAGGGATTACTTACTGACTCTGTTGGGAAGCTTT
 TTGGAAGCATTTCATGCTGAGTCCCTTTACCTTGATGTTGAAACCCATCTGGTATC
 GCTGGATCAACAACCGCCTGTGGCACACATGGCTCACCTACCTGTGGCATTATTGGAGACCA
 TGTTGGTGTAAAAGTGATTATAACTGGGGATGCATTGTCCTGGAGAAAGAAGTGTCACTA
 TCATGAACCATCGGACAAGAACATGGACTGGATGTTCTGTGAAATTGCCTGATGCGATATAGCT
 ACCTCAGATTGGAGAAAATTGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTTGGG
 CCATGCAGGCTGCTGCCTATATCTTCATTAGGAAATGGAAGGATGACAAGAGCCATTGCG
 AAGACATGATTGATTACTTTGTGATATTACGAACCACTCAACTCCTCATATTCCCAGAAG
 GGACTGATCTCACAGAAAACAGCAAGTCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTC
 AGAAAATATGAATATGTTTACATCCAAGAACTACAGGCTTACTTTGTGGTAGACCGTCTAA
 GAGAAGGTAAGAACCTTGATGCTGTCCATGATATCACTGTGGCGTATCCTCACAAACATTCTC
 AATCAGAGAACCCCTCCCCACATCCAAGGAGGACCTCAACTCTGGTGCACAAACCGTGGG
 ATCCAATAGACACCCTCCCCACATCCAAGGAGGACCTCAACTCTGGTGCACAAACCGTGGG
 AAGAGAAAAGAGAGAGGCTGCGTTCTTCTATCAAGGGAGAAGAATTTTATTACCGGAC
 AGAGTGTCACTCCACCTTGCAAGTCTGAACCTAGGGCTTGTGGTCAAATTGCTCTATAC
 TGTATTGGACCTGTTCAAGCCCTGCAATGTGCCTACTCATATATTGTACAGTCTGTAAAGT
 GGTATTTTATAATCACCATTGTAATCTTGTGCTGCAAGAGAGAATATTGGTAGACTGGAGA
 TCATAGAACTGCATGTTACCGACTTTACACAAACAGCCACATTAAATTCAAAGAAAAATG
AGTAAGATTATAAGGTTGCCATGTGAAAACCTAGAGCATATTGGAAATGTTCTAAACCTT
 TCTAAGCTCAGATGCATTTGCATGACTATGTCGAATATTCTTACTGCCATCATTATTGT
 TAAAGATATTGCACTTAATTGTGGAAAAATTGCTACAATTTTTTAATCTCTGAA
 TGTAATTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTGGGCCAGAATA
 TTATTAAACAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFGIMVSWKGIFYFILTLFWGSFFGSIFML
SPFLPLMFVNPSWYRWINNRLVATWLTPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTR
MDWMFLWNCLMRYSYLRLIKEKICLKASLGPGFWAMQAAAYIFIHRWKDDKSHFEDMIDYF
CDIHEPLQLLIFPEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTTGFVFVVDRLREGKNLD
AVHDITVAYPHNIPQSEKHLQLQDFPREIHFHRYPIDTLPTSKEDLQLWCHKRWEKEERL
RSFYQGEKNFYFTGQSVIPPCKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITI
VIFVLQERIFGGLEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGA
 CCTCCAAATCATCCATCCACCCCTGCTGTCACTGTTCTAGTGTGAGATCAACCCACAGG
 AATATCC**ATG**GCTTTGTGCTCATTGGTTCTCAGTTCTACGAGCTGGTGTCAAGGACAGTG
 GCAAGTCACTGGACCGGGCAAGTTGTCCAGGCCTGGTGGGGAGGACGCCGTCTCCTG
 CTCCCTCTTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCAGGTTCTCAGGAATCAGTCCA
 TGCTGTGGTCCACCTCTACAGAGATGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCG
 AGGGAGAACTGAGTTGTGAAGGACTCCATTGCAGGGGGCGTGTCTCTAAGGCTAAAAAA
 CATCACTCCCTCGGACATCGGCCTGTATGGGTGCTGGTTCAGTCCCAGATTACGATGAGGA
 GGCCACCTGGGAGCTGGCAGCACTGGCTCACTCCTCTATTCCATCGTGGGATA
 TGTTGACGGAGGTATCCAGTTACTCTGCCTGTCCCTCAGGCTGGTCCCCAGCCCACAGCCAA
 GTGGAAAGGTCCACAAGGACAGGATTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAG
 CCTGTATGATGTGGAGATCTCATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCAT
 CCACCTTGCTGAGCAGAGTCATGAGGTGGAATCCAAGGTATTGATAGGAGAGACGTTTCCA
 GCCCTCACCTTGGCGCCTGGCTTCTATTACTCGGGTTACTCTGTGGTGCCCTGTGGTGT
 TGTATGGGATGATAATTGTTCTCAAATCCAAAGGGAAAATCCAGGCGGAACGGACTG
 GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCGGAACACGCAGTGGAGGTGACTCT
 GGATCCAGAGACGGCTCACCGAAGCTCTGCCTTGATCTGAAAATGTAACCCATAGAAA
 AGCTCCCCAGGAGGTGCCTCACTTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTTCTCA
 GGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGACGTGGACAAAATGTAGGGTGGTATGT
 GGGAGTGTGTCGGATGACGTAGACAGGGGAAGAACAAATGTGACTTGTCTCCAAACAATGG
 GTATTGGGTCTCAGACTGACAACAGAACATTGTATTCAATTCAATCCCCATTTCAG
 CCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGTCTCCTGGACTATGAGGGGGACCAT
 CTCCTTCTCAATACAAATGACCAGTCCCTTATTATACCCCTGCTGACATGTCAGTTGAAGG
 CTTGTTGAGACCCCTATATCCAGCATGCGATGTATGACGAGGAAAAGGGACTCCATATTCA
 ATGTCCAGTGTCTGGGAT**GA**GACAGAGAAGACCCCTGCTTAAAGGGCCCCACACCACAGACC
 CAGACACAGCCAAGGGAGAGTGCTCCGACAGGTGGCCCCAGCTCCTCTCCGGAGCCTGCGC
 ACAGAGAGTCACGCCCAACTCTCCTTAGGGAGCTGAGGTTCTCTGCCCTGAGCCCTGCA
 GCAGCGGCAGTCACAGCTCCAGATGAGGGGGATTGGCCTGACCTGTGGAGTCAGAACGCC
 ATGGCTGCCCTGAAGTGGGACGGAATAGACTCACATTAGGTTAGTTGTGAAAATCCATC
 CAGCTAACGATCTTGAACAAAGTCACAAACCTCCCAGGCTCCTCATTGCTAGTCACGGACAGT
 GATTCCCTGCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTTGA
 GGGCACAGTGTGCTAATGATGTGTTTATATTACATTTCACCAACTCTGTT
 GCTTATTCCACATTAATTACTTTCTATACCAAATCACCCATGGAATAGTTATTGAACAC
 CTGCTTGTGAGGCTCAAAGAATAAGAGGAGGTAGGATTTCACTGATTCTATAAGCCCAG
 CATTACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAATACAGGTCCA
 TATCCCTCATTAACACAGACACAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAAT
 ATATTAAAGATGATATATAACTACTCAGTGTGGTTGTCCACAAATGCAGAGTTGGTTAA
 TATTAAATATCAACCAGTGTAAATTCAAGCACAATTAAAGTAAAAAGAAAACCATAAAAAAA
 AAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFKVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAV
VHLYRDGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEAT
WELRVAALGSIPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLY
DVEISIIVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVM
GMIIVFFKSKGKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAP
QEVPHSEKRFTRKSVVVASQGFQAGRHYWEVDVGQNNGWYVGVCRDDVDRGKNNVTLSNNGYW
VRLLTTEHLYFTFNPHFISLPPSTPPTRGVFLDYEGGTISFFNTNDQSLIYTLTCQFEGLL
R PYIQHAMYDEEKGTPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGA
 GAGGGCTAG
 GAAAAGAGTTGTTGGGAACCCTGGGTTATCGGCCTCGTCATCTCATATCCCTGATTGTCCT
 GGCAGTGTGCATTGGACTCACTGTTCAATTATGTGAGATATAATCAAAGAAC
 ACTACAATTA
 CTATAGCACATTGTCATTACAACACTGACAAACTATATGCTGAGTTGGCAGAGAGGCTTCTAA
 CAATTTACAGAAATGAGCCAGAGACTTGAATCAATGGTAAAAAATGCATT
 TATAAATCTCC
 ATTAAGGGAAGAATTGTCAAGTCTCAGGTTATCAAGTTCAACAGAAC
 GATGGAGTGTT
 GGCTCATATGCTGTTGATTGTAGATTCACTCTACTGAGGATCCTGAAACTGTAGATA
 AAAAT
 TGTTCAACTTGT
 TTACATGAAAAGCTGCAAGATGCTGAGGACCCCTAAAGTAGATCCTCA
 CTCAGTTAAAATTAAAAAAATCAACAAGAACAGAACAGACTATCTAAACCATTGCTGCGG
 AACACGAAGAAGTAAA
 ACTCTAGGTCAGAGTCAGGATCGTTGGTGGACAGAAGTAGAAGA
 GGGTGAATGGCCCTGGCAGGCTAGCCTGCAGTGGATGGAGTCATCGCTGTGGAGCAACCTT
 AATTAATGCCACATGGCTTGTGAGTGCTGCTCACTGTTACAACATATAAGAAC
 CCTGCCAG
 ATGGACTGCTCCTTGGAGTAACAATAAAACCTCGAAAATGAAACGGGGTCTCCGGAGAAT
 AATTGTCCATGAAAATACAAACACCCATCACATGACTATGATATTCTCTTGAGCAGCTTC
 TAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTGTCTCCCTGATGCATCCTATGAGTT
 TCAACCAGGTGATGTGATGTTGTGACAGGATTGGAGCACTGAAAATGATGGTTACAGTCA
 AAATCATCTCGACAAGCACAGGTGACTCTCATAGACGCTACAAC
 TTGCAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTAGAAGGAAA
 ACAGATGC
 ATGCCAGGGTGA
 CTGGAGGACCACTGGTTAGTCAGATGCTAGAGATATCTGGTACCTTGC
 TGGAATAGTGAGCTGGGAGATGAATGTGCGAAACCAACAAGCCTGGT
 GTTATACTAGAGT
 TACGGCCTTGC
 GGGACTGGATTACTTC
 AAAACTGGTATCTAAGAGACAAAAGCCTCATGGAA
 CAGATAACATTTTTGTTGGGTGTGGAGGCCATT
 TTAGAGATACAGAATTGGAGA
 AGACTTGCAAAACAGCTAGATTGACTGATCTCAATAAACTGTTGCTGATGCATGT
 ATT
 CTTCCCAGCTCTGTTCCGCACGTAAGC
 ATCCTGCTGCCAGATCAACTCTGT
 CATCTGTGA
 GCAATAGTTGAAACTTATGTACATAGAGAAATAGATAATACAATATTACATTACAGCCTGTA
 TTCATTGTTCTAGAAGTTGTCAGAATT
 TGACTGTTGACATAAATTGTAATGCATA
 TATACAATTGAA
 AGCAACTCCTTCTCAGTT
 CCTCAGCTCCTCTCATT
 CAGCAAATATCCA
 TTTCAAGGTGCAGAAC
 AAGGAGTGAAAGAAA
 ATATAAGAAGAAAAA
 ATCCCCTACATT
 TTGGCACAGAAA
 AGTATTAGGTGTT
 CTTAGTGGAA
 ATTAGAAATGAT
 CATATT
 CATT
 GAAAGGTCAAGCAA
 AAGACAGCAGAAC
 ACCAATCA
 CTT
 CAG
 ACT
 CTT
 CAG
 CAA
 ACA
 ACT
 ACT
 ATG
 ATAAATGTGAAGAAGATTCTGTT
 TTGACCTATAATT
 AATACAAACTTCATGCAATG
 TACTTGTTCTAAGCAA
 ATTAAAGCAA
 ATT
 TAACAA
 ATAAT
 AAC
 ACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDK
LYAEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSSQQKHGVLAHMLLICRFH
STEDPETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQS
LRIVGGTEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIK
PSKMKGRLRRITVHEKYKHPSHDYDISLAELSSPVPTNAVHRVCLPDASYEFQPGDVMFVTG
FGALKNDGYSQNHLRQAQVTLIDATTNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLV
SSDARDIWYLAGIVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAACGCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCCGAAGAACGTTCCC
 TGCCCCGATGAGCCCCCGCCGTGCGTCCCCGACTATCCCCAGGCCGGCTGGGCACCGGGCC
 CAGCGCCGACGATCGCTGCCGTTGCCCTGGGAGTAGGATGTGGTAAAGGATGGGGCTTC
 TCCCTTACGGGGCTCACA**T**GGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTAC
 GCCCTCAATCTGCTCTTGTTAATGTCCATCAGTGTGGCAGTTCTGCTTGGATGAGG
 GACTACCTAAATAATGTTCTCACTTAACAGCAGAAACGAGGGTAGAGGAAGCAGTCATTG
 ACTTACTTCCTGTGGTCATCCGGTCATGATTGCTGTTGCTGTTCTTATCATTGTGGGG
 ATGTTAGGATATTGTGGAACGGTAAAAGAAATCTGTTGCTCTGCATGGTACTTGGAAAGT
 TTGCTTGTCACTTCTGTGTAGAACTGGCTTGCGTTGGACATATGAACAGGAACATTATG
 GTTCCAGTACAATGGTCAGATATGGTCACTTGAAAGCCAGGATGACAAATTATGGATTACCT
 AGATATCGGTGGCTTACTCATGCTTGAATTTCAGAGAGAGTTAAGTGTGTGGAGTA
 GTATATTCACTGACTGGTGGAAATGACAGAGATGGACTGGCCCCAGATTCTGCTGTGTT
 AGAGAATTCCCAGGATGTTCAAACAGGCCACCAGGAAGATCTCAGTGACCTTATCAAGAG
 GGTTGTGGGAAGAAAATGTATTCTTTGAGAGGAACCAAACACTGCAGGTGCTGAGGTT
 CTGGGAATCTCATTGGGGTACACAAATCCTGGCCATGATTCTCACCATTACTCTGCTCTGG
 GCTCTGTATTATGATAGAAGGGAGCCTGGACAGACAAATGATGTCCTGAAGAATGACAAC
 TCTCAGCACCTGTCACTGTCAGTAGAACTGTTGAAACCAAGCCTGTCAAGAATCTTGAA
 CACACATCCATGGCAAACAGCTTAATACACACTTGAGATGGAGGAGTTTAAAAGAAATG
 TCACAGAAGAAAACCACAAACTTGTGTTATTGGACTTGTGAATTGAGTACATACTATGTG
 TTTCAGAAATATGTAGAAATAAAATGTTGCCATAAAATAACACCTAACGATATACTATTCTA
 TGCTTAAAATGAGGATGGAAAAGTTCATGTCATAAGTCACCACCTGGACAATAATTGATGC
 CCTTAAAATGCTGAAGACAGATGTCATAACCACTGTGTAGCCTGTATGACTTTACTGAAC
 ACAGTTATGTTGAGGCAGCATGGTTGATTAGCATTCCGCATCCATGCAAACGAGTCACA
 TATGGTGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAACTAGTATATAAGTACT
 AATTAAATGCTAACATAGGAAGTTAGAAAATACTAACATTACTCAGCGATCTATT
 TTCTGATGCTAAATAATTATATCAGAAAACCTTCAATATTGGTACTACCTAAATGTGAT
 TTTGCTGGTTACTAAAATATTCTTACCACTTAAAGAGCAAGCTAACACATTGTCTTAAGCT
 GATCAGGGATTGTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAAGTCAGTCAGTTCT
 GATAATGTTAGAATAACCATTATGAAAAGGAAAATTGTCTGTATAGCATCATTATTTTA
 GCCTTCCTGTTAACAAAGCTTACTATTCTGCTGGCTTATATTACACATATAACTGTTA
 TTTAAATACCTAACCACTAATTGAAAATTACCACTGTGATACATAGGAATCATTATTCTAGA
 ATGTTAGTCTGGCTTTAGGAAGTATTAATAAGAAAATTGCACATAACTTAGTGATTCAAGAA
 AGGACTTGTATGCTGTTCTCCAAACAAAGCAATAGTCTCCAAGTCATATAAAATTCTACAGA
 AAATAGTGTCTTTCTCCAGAAAATGCTGTGAGAATCATTAAACATGTGACAATTAG
 AGATTCTTGTGTTATTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATT
 TTTACAAGAGTATAGTATATTGAAATGGAAAAGTGCATTACTGTATTGTTGTAT
 TTTGTTATTCTCAGAACATATGAAAGAAAATTAAATGTGTCAATAAAATTCTAGAGAG
 TAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAWSAWMRDYLNNVLTLAETRVEEAVILTYFPVV
HPVMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWS
DMVTLKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMDWPPDSCCVREFPGC
SKQAHQEDLSLDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDR
REPDTDQMMMSIKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACACTCATCCTCATCCTCTTCCTCTGATAAAAGCCC
 CTACCAAGTGCTGATAAAGTCTTCTCGTGAGAGCCTAGAGGCCTAAAAAAAAGTGCTTGA
 AAGAGAAGGGACAAAGGAACACCAAGTATTAAGAGGATTTCCAGTGTCTGGCAGTTGGTC
 CAGAAGGATGCCTCCATTCTGCTTCTCACCTGCCTCTCATCACAGGCACCTCCGTGTCAACC
 CGTGGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGCCCTGGAGGAACACTGACCA
 CCAGTTGGATGAGTCTCAAGGTCTCCTCTATGTGACAACCAGTGAATGGGGAGTGGTACCA
 CTTCACGGGCATGGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCAGTGTGGAAC
 CCACGCACCTGTCTGGCTCAATGGCAGCCACCCCCCTAGAAGGCGACGGCATTGTGCAACGCCA
 GGCTTGTGCCAGCTCAATGGGAACGTGCTCTGGAACACCACGGTGGAAAGTCAAGGCTTG
 CCCTGGAGGCTACTATGTATCGTCTGACCAAGGCCAGCGTCTGCTTCCACGTCTACTGTGG
 TCATTTTATGACATCTGCAGCAGGACTGCCATGGCAGCTGCTCAGATAACCAGCGAGTGCAC
 ATGCGCTCCAGGAACGTGCTAGGCCCTGACAGGCAGACATGCTTGTGAAAGAATGTGA
 GCAAAACAACGGTGGCTGCAGCCACTCTGCCTGGATCTGAGATCTGTGAAACCTCAAAACTCCTACCGCTGTGAGTG
 TGGGGTTGGCGTGTGCTAAGAAGTGAAGACTTGTGAAGACGTTGAAGGATGCCACAA
 TAACAATGGTGGCTGCAGCCACTCTGCCTGGATCTGAGAAAGGCTACCAAGTGTGAATGTCC
 CCGGGGCTGGTGTCTGAGGATAACCACACTGCCAAGTCCCTGTGTTGGCCTGGAGCTTCTGACCAACAC
 TGCCATTGAAGTGAACATCCCCAGGGAGCTGGTGGTGGCCTGGAGCTTCTGACCAACAC
 CTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACATCCTCTCTCAAGACATGTGG
 TACAGTGGTCGATGTGGTGAATGACAAGATTGTGGCAGCAACCTCGTACAGGTCTACCCAA
 GCAGACCCCCGGGAGCAGCGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCGGTGAC
 CTGCGAGTTCCACGCCTGTACACCATTCTGAAGGATACGTTCCAACCTTCGAAACTCC
 ACTGGAAATCATGAGCCGAAATCATGGGATCTCCCATTCACTCTGGAGATCTCAAGGACAA
 TGAGTTGAAGAGCCTTACCGGGAGCTCTGCCACCCCTCAAGCTCGTACTCCCTACTT
 TGGCATTGAGCCC GTGGTGCACGTGAGCGGCTGGAAAGCTTGGGAGAGCTGCTTGCAC
 CCCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCATCCGGATGGCTGTGTTCAGA
 TGACTCGTAAAGCAGTACACATCCCAGGATCACCTAGCAAAGCAGCTCCAGGTCCCTGTCTT
 CAAGTTGTGGCAAAGACCAACAAGGAAGTGTCTGCACTGCCGGTTCTGTGTGGAGT
 GTTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGGGAATGCGTCGTGGGGCAGGAGG
 AGAGGACTCAGCCGGTCTACAGGGCCAGACGCTAACAGGCAGGCCATCGCATCGACTGGGA
 GGACTTAGTCGTAGCCATACTCGAGTCCCTGCATTGGACGGCTTGCTCTTGGAGCTTCTC
 CCCCCACCGCCCTTAAGAACATCTGCCAACAGCTGGGTTAGACTTCACACTGTGAGTTCA
 ACTCCCAGCACCAACTCACTCTGATTCTGGTCATTCACTGGGACAGGTACAGCACTGCTG
 AACAAATGTGGCCTGGGTGGGTTCTACCTTCTAGGGTTGAAAACCTAAACTGTCCACCCAGAA
 AGACACTCACCCATTCCCTCATTTCTTACACTTAAACACCTCGTGTATGGTGCACATC
 AGACCACAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCCCTAGAAAAATTAAACAG
 TTACTGAAATTATGACTTAAATACCCAATGACTCCTTAAATATGTAAATTATAGTTACCTT
 GAAATTCAATTCAAATGCAGACTAATTATAGGAAATTGGAAGTGTATCAATAACAGTAT
 ATAATT

FIGURE 110

MPPFLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFT
GMAGDAMPTFCIPEHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPG
GYYVYRLTKPSVCFHVYCGHFYDICDEDCHGSCSDTSECTCAPTVLGPDRQTCFDENECEQN
NGGCSEICVNLKNSYRCECGVGRVLRSRGKTCEDVEGCHNNNGCASHSCLGSEKGYQCECPRG
LVLSEDNHTCQVPVLCKSNAIEVNIPRELVGGLEFLFLTNTSCRGVSNGTHVNILFSLKTCGT
VDVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVTCEFPRLYTISEGYVPNLRNSPLE
IMSRNHGIFPFTLEIFKDNEFEPYREALPTLKLRLDSLYFGIEPVVHVSGLESLVESCFA
SKIDEVLKYLIIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRVLVCVLD
ERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTGCTCAGCGGACAAGGATGCTGGCGTGAGGGACCAAGGCCTGCCCT
 GCACTCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGCCAGCCAGGACC
 TGTGTGGGGAGGCCCTCCTGCTGCCTTGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACC
 GGGAGGATCACAGAGCCAGC**ATG**TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCG
 ATGTCAAACCCCTGCGCAAACCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCA
 TCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGTGGTTGCCTCATCAAGGTGATTG
 TGGATAAAATACTACTTCCTCTGCCGGCAGCCTCTCCACTTCATCCGAGGAAGCAGCTGTGTG
 ACGGAGAGCTGGACTGTCCCTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTCCCCGAAG
 GCCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCA
 CAGGGAACTGGTTCTGCCTGTTGACAACCTCACAGAAGCTCGCTGAGACAGCCTGTA
 GGCAGATGGGCTACAGCAGAGCTGTGGAGATTGGCCAGACCAGGATCTGGATGTTGAAA
 TCACAGAAAACAGCCAGGAGCTCGCATCGGAACACTCAAGTGGCCCTGTCTCTCAGGCTCCC
 TGGTCTCCCTGCACTGTCTGCCTGTGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGG
 AGGAGGCCTCTGTGGATTCTGGCCTTGGCAGGTCAAGTACGACAAACAGCACGTCT
 GTGGAGGGAGCATCCTGGACCCCCACTGGTCCTCACGGCAGCCCAGTCTCAGGAAACATA
 CCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGCAGCTCCATCCCTGG
 CTGTGGCCAAGATCATCATCATTGAATTCAACCCATGTACCCAAAGACAATGACATCGCCC
 TCATGAAGCTGCAGTCCCACTCACCTTCAGGCACAGTCAGGCCATCTGTCTGCCCTCT
 TTGATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATTGGATGGGCTTACGAAGCAGA
 ATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCAGGTCAATTGACAGCACACGGT
 GCAATGCAGACGATGCGTACCAAGGGGAAGTCACCGAGAAGATGATGTGCAGGCATCCGG
 AAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCCCTGATGTACCAATCTGACCA
 GGCATGTGGTGGCATCGTTAGCTGGGCTATGGCTGCCGGCCAGCAGGAGTAT
 ACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG**TAA**TGCT
 GCTGCCCTTGCACTGCTGGAGCCGCTTCCCTGCCCTGCCACCTGGGATCCCCAA
 AGTCAGACACAGAGCAAGAGTCCCTGGTACACCCCTCTGCCACAGCCTCAGCATTCTT
 GGAGCAGCAAAGGGCTCAATTCTGTAAGAGACCCCTCGCAGCCCAGAGCGCCAGAGGAAG
 TCAGCAGCCCTAGCTCGGCCACACTGGTGCTCCCAGCATCCCAGGGAGAGACACAGCCCAG
 GAACAAGGTCTCAGGGTATTGCTAACCCAAGAAGGAACCTTCCCACACTACTGAATGGAAGC
 AGGCTGTCTGTAAAAGCCCAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCA
 GCCCTGTCCGTCTCACCCATCCCCAAGCCTACTAGAGCAAGAAACCAGTTGTAATATAAAAT
 GCACTGCCCTACTGTTGGTATGACTACCGTTACCTACTGTTGTCATTGTTATTACAGCTATGG
 CCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRI PMETFRKVGIP III ALLSLASII IVVVLIKVILDKYFL
CGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLD SATGNWFSA
CFDNFTEAL AETACRQM GYSRAVEIGPDQ DLDV VEITENS QELRMRN SSGPCLSGSLVSLHCL
ACGKSLKT PRVVG GEEASV DSWP WQVSI QYDK QHV CGGSILD PHW VLTA AHCF RKHT DVFN WK
VRAGSDKLGSFPSLAVAKIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTP
ATPLWIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDT C
QGDGGPLMYQSDQWHVVGIVSWG YGC GGPSTPGV YT KVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (type II)

FIGURE 113

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGAT
 TATAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTTAGCAACTAAAAAGCCACAG
 GAGTTGAACTGCTAGGATTCTGACT**ATG**CTGTGGTGGCTAGTGCTCCTACTCCTACCTACATT
 AAAATCTGTTTTGTTCTTGTAACTAGCCTTACCTCCTAACACACAGAGGATCTGTCACT
 GTGGCTCTGGCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTCTACCCACACCGT
 CCCCTCGAAGCCGGGGACAGCCTCACCTGCTGGCCTCTCGCTGGAGCAGTGCCCTACCAAC
 TGTCTCACGTCTGGAGGCAGTGACTCGGGCAGTGCAGGTAGCTGAGCCTTGGTAGCTGCAG
 CTTTCAAGGTGGGCCTGGCCCTGGCGTAGAAGGGAT**TGA**CAAGCCCAGAAGATTCAAGGCG
 ATGGCTCCCAGTCCCCAGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGCCAGGACGG
 GCCGTGGACACCTGCTCAGAAGCAGTGGGTGAGACATCACGCTGCCGCCATCTAACCTTT
 CATGTCCCTGCACATCACCTGATCCATGGGCTAATCTGAACCTGTCCCAAGGAACCCAGAGCT
 TGAGTGAGCTGTGGCTCAGACCCAGAAGGGGTCTGCTTAGACCACCTGGTTATGTGACAGGA
 CTTGCATTCTCCTGGAACATGAGGGAACGCCGGAGGAAAGCAAAGTGGCAGGGAAAGGAACCTG
 TGCCAAATTATGGGTAGAAAAGATGGAGGTGTTGGTTATCACAAGGCATCGAGTCTCCTGC
 ATTCACTGGACATGTGGGGAAAGGGCTGCCATGGCGATGACACACTCGGGACTCACCTCTG
 GGGCCATCAGACAGCCGTTCCGCCCGATCCACGTACAGCTGCTGAAGGGCAACTGCAGGC
 CGATGCTCTCATCAGCCAGGCAGCAGCCAAATCTGCGATCACAGCCAGGGCAGCCGTCTG
 GGAAGGAGCAAGCAAAGTGACCATTCTCCTCCCTCCTGAGAGGCCCTCCTATGT
 CCCTACTAAAGCCACCAGCAAGACATAGCTGACAGGGCTAATGGCTCAGTGTGGCCAGGA
 GGTCAAGGCCTGAGAGCTGATCAGAAGGGCTGCTGTGCGAACACGGAAATGCCTCCAGT
 AAGCACAGGCTGAAAATCCCCAGGCAAAGGACTGTGTGGCTCAATTAAATCATGTTCTAGT
 ATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTTCAAATGATCTCCAAGGCCCT
 TATAACCCAGGAGACTTGTATTGAAATTGAAACCCCAAATCCAAACCTAAGAACCCAGGTGCA
 TTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACATTGGTAAACCCCTGTCTC
 GGTAGATCACCTGAGGTCAAGGAGCTGATCAGGAGCTGAGACAGCCAGCCTGGCCAACATGGTAAACCCCTGTCTC
 TACTAAAAATACAAAAAAACTAGCCAGGCATGGTGGTGTGCCTGTATCCCAGCTACTCGGG
 AGGCTGAGACAGGAGAATTACTGAACCTGGGAGGTGAAGGAGGCTGAGACAGGAGAATCACT
 TCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTATGGTTATT
 GTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTAS
PCWPLAGAVPSPTVSRLTEALTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGC
AAAGAACCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTAAATCCAA
GAAAATATGTAAATCACTTAAGATTGTGGACTGGTGTGTTGGTATCCTGCCCTAACTCTAAT
TGTCTGTTGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGA
GCACACTTCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAG
AACTGAAATATTCAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTAAAAA
CGGATACACTGGCATCTACTCGTGGTCTTCAAAATGTTATCAAAACTCAGATTAAAGT
GATTCTGAATTTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTT
CTTGAAACAGTCAGTGATTGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTCTTAA
AAATTCCAAAATTCTGGAGATTGTGATAACGTGACCATGTATTGGATCAATCCACTCTAAT
ATCAGTTCTGAGTTACAAGACTTGAGGGAGGGAGAAGATCTCACTTCTGCCAACGA
AAAAAAAGGGATTGAACAAAATGAACAGTGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCG
TCACGCCAGACAAGCAAGTGAGGAAGAACTCCAATAATGACTATACTGAAAATGGAATAGA
ATTGATCCCAGCTGGATGAGAGAGGTTATTGTTGTATTACTGCCGTGAGGCAACCGCTA
TTGCCGCCGCTGTGAACCTTACTAGGCTACTACCCATATCCACTGCTACCAAGGAGG
ACGAGTCATCTGTCGTGTCATCATGCCTTGTAACTGGTGGTGGCCCGCATGCTGGGAGGGT
CTAAATAGGAGGTTGAGCTCAAATGCTAAACTGCTGGCAACATATAATAATGCATGCTATT
CAATGAATTCTGCCTATGAGGCATCTGGCCCTGGTAGCCAGCTCTCCAGAATTACTGTAG
GTAATTCTCTTCAATAAACTTCTACATTACCAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKSKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPKKAYD
MEHTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQI
KVIPEFSEPEEEIDENEETTTFFEQSVIWVPAEKPIENRDFLKNKILEICDNVTMYWINPT
LISVSELQDFEEEEDLHFPAKEKKGIEQNEQWVVPQVKVEKTRHARQASEEEELPINDYTENG
IEFDPMLERGYCCIYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTCACACCTCGGCAGCAGGAGGGCGGCAGCTTCAGC
 AGGCAGGGCAGGGCGGGCGGCCAGGATC**ATG**TCCACCACCATGCCAAGTGGTGGCGTTCC
 TGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCAGGATGGACATGTGGAGCACCCAGG
 ACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGA
 GGCAGAGTTCAAGGCTTCACCGAACATGCAGGCCATTTCACCATCCTGGACTTCCAGCCATGC
 TGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGCCTGGGTGCCATTGGCCTCCTGGTAT
 CCATCTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAACATGA
 CACTGACCTCCGGGATCATGTTCATTGTCTCAGGTCTTGCAATTGCTGGAGTGTCTGTGT
 TTGCCAACATGCTGGTACTAACTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTG
 GGATGGTGCAGACTGTTCAAGACCAGGTACACATTGGTGCAGCTCTGGTGGCTGGCACCAG
 CTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATGCCCTGCCGGGCCTGGCACCAG
 AAGAAACCAACTACAAAGCCGTTCTTATCATGCCCTCAGGCCACAGTGTGCCTACAAGCCTG
 GAGGCTTCAAGGCCAGCACTGGCTTGGGTCCAACACCAAAAAACAAGAAGATATCGATGGAG
 GTGCCGCACAGAGGACGAGGTACAATCTTATCCTCCAAGCACGACTATGTG**TAAT**GCTCTA
 AGACCTCTCAGCACGGCGGAAGAAACTCCCAGAGCTCACCCAAAAACAAGGAGATCCCA
 TCTAGATTCTTCTGCTTGTGACTCACAGCTGGAAAGTTAGAAAAGCCTCGATTCATCTTG
 GAGAGGCCAAATGGTCTTAGCCTCAGTCTGTCTCAAATATTCCACCAAAACAGCTGAG
 TTATTATGAATTAGAGGCTATAGCTCACATTCAATCCTCTATTCTTTAAATATAA
 CTTCTACTCTGATGAGAGAATGTGGTTTAATCTCTCTCACATTGATGATTAGACAG
 ACTCCCCCTCTCCTCTAGTCAATAAACCCATTGATGATCTATTCCCAGCTATCCCCAAG
 AAAACTTTGAAAGGAAAGAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGTCTC
 CCCACCCCCAACTGGCTAGTAATAACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATT
 TGTAATCTCTCCAGCCCATTGATCTCGGTTCTTACACTGTGATCTAAAGTTACCAAAACCA
 AAGTCATTTCAGTTGAGGCAACCAACCTTCTACTGCTGTTGACATCTTCTTATTACAGC
 AACACCATTCTAGGAGTTCTGAGCTCTCCACTGGAGTCCTTCTGCGGGTCAGAAA
 TTGTCCCTAGATGAATGAGAAAATTATTTTTAATTAAAGTCCTAAATATAGTTAAAATAA
 ATAATGTTTAGTAAATGATACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAG
 AAGGAAATGAAAAATAATTGCTTGCATTGCTATATGGTACTTGTAAAGTCATGCTTAA
 GTACAAATTCCATGAAAAGCTCACACCTGTAATCCTAGCACCTGGAGGCTGAGGAGGAAGG
 ATCACTTGAGCCCAGAAGTCAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTCACAAA
 ATACAGAGAGAAAAATCAGCCAGTCATGGTGGCATACACCTGTAGTCCCAGCATTCCGGAG
 GCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGTTGGGCTGCAGTGAGCCATGATCACACC
 ACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAATAAAATAATGGA
 ACACAGCAAGTCCTAGGAAGTAGGTTAAAACATAATTCTTAA

FIGURE 118

MSTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTEC
RPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFI
VSGLCIAIGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGG
VMMCIACRGLAPEETNYKAVSYHASGHHSVAYKPGGFCASTGFGSNTKNKKIYDGGARTEDEVQ
SYPSKHDYV

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTCTTGTGGCACAGAGAACCCCTGCTCAAAGCAGAAGTAGCAGTTCCGG
 AGTCCAGCTGGCTAAAACATCCCAGAGGATA**ATG**GCAACCCATGCCTAGAAATCGCTGGG
 CTGTTCTGGTGGTGGAAATGGTGGCACAGTGGCTGTCAGTGCATGCCTCAGTGGAGA
 GTGTCGGCCTTCATTGAAAACAACATCGTGGTTTGAAAACCTCTGGGAAGGACTGTGGATG
 AATTGCGTGAGGCAGGCTAACATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTT
 TCTCCGGACCTACAGGCAGCCAGAGGACTGATGTGTGCTGCTCCGTGATGTCCTCTGGCT
 TTCATGATGCCATCCTGGCATGAAATGCACCAGGTGCACGGGGACAATGAGAAGGTGAAG
 GCTCACATTCTGCTGACGGCTGGAATCATCTCATCACGGCATGGTGGTGCTCATCCCT
 GTGAGCTGGTTGCCAATGCCATCATCAGAGATTCTATAACTCAATAGTAATGTTGCCAA
 AACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCACTGGTGCTGATTGTTGGA
 GGAGCTCTGTTGCTGCGTTTTGCAACGAAAAGAGCAGTAGCTACAGATACTCGATA
 CCTTCCCATCGCACACCCAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCGTCTACTCC
 AGAAGTCAGTATGTG**TAG**TTGTATGTTTTAACTTAACTAAAGCCATGCAAATGACA
 AAAATCTATATTACTTCTCAAAATGGACCCAAAGAAACTTGTATTACTGTTCTTAACG
 CTAATCTTAATTACAGGAACGTGCATCAGCTATTGATTCTATAAGCTATTCAAGCAGAA
 TGAGATATTAAACCAATGCTTGATTGTTCTAGAAAGTATAGTAATTGTTCTAAGGTGG
 TTCAAGCATCTACTCTTTATCATTACTCAAAATGACATTGCTAAAGACTGCATTATT
 ACTACTGTAATTCTCCACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCA
 CATAGAGACATGCTTATGGTTATTAAATGAAATGCCAGTCCATTACACTGAATAAAAT
 AGAACTCAACTATTGCTTTCAGGGAAATCATGGTAGGGTTGAAGAAGGTTACTATTAAATTG
 TTTAAAAACAGCTAGGGATTAATGTCCTCCATTATAATGAAGATTAAAATGAAGGCTTAA
 TCAGCATTGTAAGGAAATTGAATGGCTTCTGATATGCTGTTTAGCCTAGGAGTTAGAA
 ATCCTAACTTCTTATCCTCTTCCCAGAGGCTTTCTGTATTAAATTAAACATT
 TTTAAAACGCAGATATTGTCAAGGGCTTGCATTCAAACGTCTTCCAGGGCTACTC
 AGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGTGTTAGGAAAGTGAAAATATT
 GTTTGTATTGAAGAAGAATGATGCATTGACAAGAAATCATATATGTATGGATATT
 TAATAAGTATTGAGTACAGACTTGGAGTTCATCAATATAAAAGAGCAGAAAATAT
 GTCTGGTTTCATTGCTTACCAAAAAACACAACAAAAAGTTGTCTTGAGAAGCTTC
 ACCTGCTCCTATGTGGTACCTGAGTCAAAATTGTCATTGTTCTGTGAAAAATAATT
 CTTCTGTACCATTCTGTTAGTTACTAAAATCTGAAATACTGTATTCTGTTATT
 CCAAATTGATGAAACTGACAATCCAATTGAAAGTTGTGTCGACGTCTGTCTAGCTAAAT
 GAATGTGTTCTATTGCTTATACATTATTAATAAAATTGTACATTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTAVTVMPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQ
CKIYDSLLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIF
IITGMVVLIPVSWVANAIIRDIFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCN
EKSSSYRYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAGGCGCATTGATGCAGCCTGC GGCGGCCTCGGAGCGCGGGCG
AGCCAGACGCTGACCACGTT CCTCTCGGTCTCCTCCGCCTCCAGCTCCGCCTGCCCGC
AGCCGGGAGCC**ATG**CGACCCCAGGGCCCCGCCCTCCCCGCAGCGGCTCCGCCTCCTGC
TGCTCCTGCTGCTGCAGCTGCCCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGAAAGCAA
AGGCGCAGCTCCGGCAGAGGGAGGTGGACCTGTATAATGGAATGTGCTTACAAGGCCAG
CAGGAGTGCCTGGTCGAGACGGGAGCCCTGGGCCAATGTTATTCCGGTACACCTGGATCC
CAGGTGGGATGGATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTGAGGAGTCCT
GGACACCCA ACTACAAGCAGTGTTCATGGAGTTCAATTGCAATTGGCATAGATCTGGAAAA
TTGCGGAGTGTACATTACAAAGATGCGTTCAAATAGTGCTCTAACAGAGTTGTCAGTGGCT
CACTTCGGCTAAAATGCAGAAATGCATGCTGTCAGCGTTGGTATTCACATTCAATGGAGCTG
AATGTTCAAGGACCTCTCCATTGAAGCTATAATTATTGGACCAAGGAAGCCCTGAAATGA
ATTCAACAATTAATATTCACTGCACTTCTGTGGAAGGACTTGTGAAGGAATTGGTGCTG
GATTAGTGGATGTTGCTATCTGGGTTGGCACTGTTCAAGATTACCAAAAGGAGATGCTTCTA
CTGGATGGAATTCAAGTTCTCGCATCATTATTGAAGAACTACCAAA**TAA**ATGCTTAATT
CATTTGCTACCTCTTTTATTATGCCTTGGAATGGTCACTTAAATGACATTAAATAAG
TTTATGTATACATCTGAATGAAAAGCAAAGCTAAATATGTTACAGACCAAAGTGTGATTCA
CACTGTTTAAATCTAGCATTATTCACTTGTCAATCAAAAGTGGTTCAATATTTTT
TAGTTGGTTAGAATACTTCTCATAGTCACATTCTCAACCTATAATTGGAATTGTTG
TGGTCTTTGTTCTTAGTATAGCATTAAAAAATATAAAAGCTACCAATTCTT
TACAATTGTAAATGTTAAGAATTCTTATCTGTTAAATAAAATTATTCCAACA

FIGURE 122

MRPQGPAAS PQRLRGLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVP
GRDGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAEC
TFTKMRNSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGPLPIEAIYLDQGSPEMNSTI
NIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTCTGGGCTCCAACGCAGCTCTGTGGCTG
 AACTGGGTGCTCATCACGGGAACGTGCTGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCC
 CCAAATTGCCTGGAAGAATAACATCATGTTTCGATAAGAAGAAATTGTAGGATCCAGTTTT
 TTTTAACCGCCCTCCCCACCCCCAAAAAAACTGTAAAGATGCAAAACGTAATATCCAT
 GAAGATCCTATTACCTAGGAAGATTGATGTTGCTGCGAATGCGGTGTTGGGATTATT
 GTTCTGGAGTGTCTGCGTGGCTGGCAAAGAATAATGTTCAAATCGGTCCATCTCCAAG
 GGGTCCAATTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGG
 CTGTCATGCAACTGGCCCCTAACGCAAAGCAGCTAACGGACGACCTTGAACAATACAA
 AGGATGGGTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCC
 ACTGTCTTACTGACAATGCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAA
 GCAAATGGTATATTGTAATCTCAGAAATTACAGGAGATACCCTCAAGTATACTGCTGGT
 TGCTTAGGTTGTCCTCGCTATAACAGCCTCAAAAACCTAACGTTAACGAAATTCAATTAAAGGG
 CTCAACCAGCTCACCTGGCTATAACCTTGACCATAACCATAACGAAATTGACGAAAATGCT
 TTTAATGGAATACGCAGACTCAAAGAGCTGATTCTTAGTCCAATAGAATCTCCTATTTCTT
 AACAAATACCTTCAGACCTGTGACAAATTACGGAACCTGGATCTGCTCTATAATCAGCTGCAT
 TCTCTGGATCTGAACAGTTCGGGCTTGCAGCTGAGTTACATTACGGTCTAAC
 TCCCTGAGAACCATCCCTGTGCGAATATTCCAAGACTGCCAACCTGGAACCTTGGACCTG
 GGATATAACCGGATCGAAGTTAGCCAGGAATGTCTTGCTGGCATGATCAGACTCAAAGAA
 CTTCACCTGGAGCACAATCAATTTCAGCTCAACCTGGCCCTTTCCAAGGTTGGTCAGC
 CTTCAGAACCTTACTTGAGTGGAAATAAAATCAGTGTCAAGGACAGACATGTCCTGGACC
 TGGAGCTCCTACAAAGGCTTGATTATCAGGCAATGAGATCGAAGCTTCAGTGGACCCAGT
 GTTTCCAGTGTGTCCGAATCTGAGCGCCTAACCTGGATTCCAACAGCTCACATTATT
 GGTCAAGAGATTGGATTCTGGATATCCCTCAATGACATCAGTCTGCTGGAAATATATGG
 GAATGCAGCAGAAATTGCTCCCTGTAAACTGGCTGAAAAGTTAAAGGTCTAAGGGAG
 AATACAATTATCTGTGCCAGTCCAAAGAGCTGCAAGGAGTAAATGTGATCGATGCAGTGAAG
 AACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTGATCTGCCAGGGCTCTCCCAAAG
 CCGACGTTAACGCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTTGGCCCCGACG
 GTGGGAGCCACAGAGCCGCCAGAGACCGATGCTGACGCCAGCACATCTTTCCATAAAA
 ATCATCGCGGGCAGCGTGGCGTTTCTGTCCGTGCTCGTCATCCTGCTGGTTATCTACGTG
 TCATGGAAGCGGTACCCCTGCGAGCATGAAGCAGCTGCAGCGCTCCCTCATGCGAAGGCAC
 AGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTTATGTAGAT
 TATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGACCCCTGCACCTAT
 AACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTGTGATAAAAAGAGCTCTAAAGC
 TGGGAAATAAGTGGTGCTTATTGAACCTGGTGACTATCAAGGGAACGCGATGCCCTTC
 CCCTCCCTCTCCCTCACTTGGTGGCAAGATCCTCCTGTCCGTTAGTGCATTATA
 ATACTGGTCATTTCTCTCATACATAATCAACCCATTGAAATTAAATACCACAATCAATGT
 GAAGCTTGAACCTCCGGTTAATATAATACCTATTGTATAAGACCCTTACTGATTCCATTAAT
 GTCGCATTGTTAAGATAAAAACCTTCTCATAGGTAAAAAAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGC
LGLSLRYNSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLN
NTFRPVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHRSNSLRTIPVRIFQDCRNLELLDG
YNRIRSLARNVFAGMIRLKELHLEHNQFSKLNLAFLPRLVSLQNLQWNKISVIGQTMSWTW
SSLQRLDLSGNEIEAFSGPSVFQCVPNLQRLNLDNSNKLTFIGQEILDWSISLNDISLAGNIWE
CSRNICSLVNWLKSFKGLRENTIICASPKELOGVNVIDAVKNYSICGKSTTERFDLARALPKP
TFKPKLPRPKHESKPPLPPTVGATEPGPETDADAEHISFHKI IAGSVALFLSVLVILLVIYVS
WKRYPASMKQLQQRSLMRRHRKKKRQSLKQMTPSTQE FYVDYKPTNTETSEMLLNGTGPCTYN
KSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCCTACTGCTGA**ATG**TCCGTCCGGAGGAGGAGGAGGCTTGTGCCGC
TGACCCAGAGATGGCCCCGAGCGAGCAAATTCTACTGTCCGGCTGCGCGCTACCGTGGCCG
AGCTAGCAACCTTCCCCTGGATCTCACAAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTC
TTGCTCGGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAAGGGAAATGGTGCGCACAGCCC
TAGGGATCATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCCATTACA
GACACGTAGTGTATTCTGGAGGTGCAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTG
GCAAAAGTGAAGATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTG
TTATTGCCAGTTTAGCCAATCCAAC TGACCTAGTGAAGGTTAGATGCAAATGGAAGGAA
AAAGGAAACTGGAAGGAAAACCATTGCGATT CGTGGTGTACATCATGCATTGCAAATCT
TAGCTGAAGGAGGAATACGAGGGCTTGGCAGGCTGGTACCCAATATAACAAAGAGCAGCAC
TGGTGAATATGGGAGATTAAACCACTTATGATACTGAAACACTACTTGGTATTGAATACAC
CACTTGAGGACAATATCATGACTCACGGTTATCAAGTTATGTTCTGGACTGGTAGCTTCTA
TTCTGGAACACCAGCGATGTCAAAAGCAGAATAATGAATCAACCACGAGATAAACAG
GAAGGGACTTTGTATAAATCATCGACTGACTGCTTGATTCAAGGTGAAGGAT
TCATGAGTCTATATAAAGGCTTTACCATCTGGCTGAGAATGACCCCTGGTCAATGGT
TCTGGCTTACTTATGAAAAAATCAGAGAGATGAGTGGAGTCAGTCCATT**TAA**

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGAR
ESAPYRGMVRTALGIIEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLREVVFGKSEDEHYP
LWKSVIGGMMAGVIGQFLANPTDLVKVQMQMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGL
WAGWVPNIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVI
KSrimnQPRDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIR
EMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCAAGCAGGTGGCGGCCGGCAGGAGAGCAGGCCGGCGTCAGCTCCTCG
 ACCCCCCGTGTGGGCTAGTCCAGCGAGGCAGGGCGCTGGCGCTGGTACGGGGGCCTCGGGGGCATCGGCGCGGC
 ATGGAGCGGTGGCGACCGGCTGGCGCTGGTACGGGGGCCTCGGGGGCATCGGCGCGGC
 GTGGCCCGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGCTGCGCCGCAGTGTGGCAAC
 ATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCGGACTTGATCCCCTACAGA
 TGTGACCTATCAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGC
 GGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCGGCCTGACACCCTGCTCTCAGGCAGC
 ACCAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGAA
 GCCTACCAGTCCATGAAGGAGCGGAATGTGGACGATGGCACATCATTAACATCAATAGCATG
 TCTGGCCACCGAGTGTACCCCTGTCTGTGACCCACTCTATAGTGCCACCAAGTATGCCGTC
 ACTGCGCTGACAGAGGGACTGAGGCAAGAGCTCGGGAGGCCAGACCCACATCCGAGCCACG
 TGCATCTCTCCAGGTGTGGAGACACAATTGCCTCAAACCTCCACGACAAGGACCCCTGAG
 AAGGCAGCTGCCACCTATGAGCAAATGAAGTGTCTCAAACCGAGGATGTGGCCGAGGCTGTT
 ATCTACGTCCTCAGCACCCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACGGAG
 CAGGTGACCTAGTGACTGTGGAGCTCCTCCTCCCTCCACCCCTCATGGCTGCCTCCTG
 CCTCTGGATTTAGGTGTTGATTCTGGATCACGGGATACCACCTCCTGTCCACACCCGACC
 AGGGCTAGAAAATTGTTGAGATTATATCATCTTGTCAAATTGCTTCAGTTGTAAATG
 TGAAAAATGGCTGGGAAAGGAGGTGGTGTCCCTAATTGTTACTTGTAACTTGTCTTG
 TGCCCCCTGGCACTTGGCCTTGTCTGCTCTCAGTGTCTCCCTTGACATGGAAAGGAGTT
 GTGGCCAAAATCCCCATCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGTGGCAGAGG
 GAGGCCTCACCTATATCTGTGTTATCCAGGGCTCCAGACTCCTCCTCTGCCTGCC
 ACTGCACCCCTCTCCCCCTATCTATCTCCTCTGGCTCCCCAGCCCAGTCTGGCTTGT
 CCCCTCCTGGGTCACTCCACTCTGACTATGGCAGCAGAACACCAGGGCCTGGC
 CCAGTGGATTCATGGTGATCATTAAAAAGAAAAATCGCAACCAAAAAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGYPGT
LIPYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVNVLALS
ICTREAYQSMKERNVDDGHIININSMSGHRVLPLSVTHFYSATKYAVTALTEGLRQELREAQT
HIRATCISPGVVETQFAFKLHDKDPEKAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQ
MRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AAC TTCTAC **ATG**GGCCTCCTGCTGCTGGTGCTCTCCTCAGCCTCCTGCCGGTGGCCTACACC
ATCATGTCCCTCCCACCCCTCCTTGACTGCAGGCCGTTCAAGGTGCAGAGTCTCAGTTGCCCGG
GAGCACCTCCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCAGAATTCCAGTTCTGGTT
TCATGCCAGCCTGTAAAAGGCCATGGAACCTTGGGTGAATCACCGATGCCATTAAAGAGGGTT
TTCTGCCAGGATGGAAATGTTAGGTGTTCTGTGTCTGCGCTGTTCAATTTCAGTAGGCCACCAG
CCACCTGTGGCCGTTGAGTGCTTGAAA**TGA**GGAACTGAGAAAATTAAATTTCATGTATTTT
CTCATTATTATTAAATTAACTGATAGTTGTACATATTGGGGTACATGTGATATTGG
ATACATGTATAACAATATAATGATCAAATCAGGGTAACGGGATATCCATCACATCAAACAT
TTATTTTTATTCTTTAGACAGAGTCTCACTCTGTACCCAGGCTGGAGTGCAGTGGTGC
ATCTCAGCTTACTGCAACCTCTGCCCTGCCAGGTTCAAGCGATTCTCATGCCCTCACCTCCAA
GTAGCTGGACTACAGGCATGCACCACAATGCCCAACTAATTGTATTAGAGACG
GGGTTTGCCATGTTGCCAGGCTGGCCTGAACCTCTGGCCTCAAACAATCCACTGCCTCG
GCCTCCAAAGTGTATGATTACAGGCGTGAGCCACCGTGCCTGGCTAAACATTATCTTT
CTTGTTGGAACTTGAAATTATAACATGAATTATTGTTAACTGTCATCTCCCTGCTGTG
CTATGGAACACTGGACTTCTCCCTCTATCTAACGTATATTGTACCAAGTTAACCAACCGT
ACTTCATCCCCACTCCTCTATCCTCCAACCTCTGATCACCTCATTCTACTCTACCTC
CATGAGATCCACTTTAGCTCCCACATGTGAGTAAGAAAATGCAATTGTCTTCTGTG
CCTGGCTTATTCACTAACATAATGACTCCTGTTCCATGTTGCTGCAAATGACAGGA
TTTCGTTCTTAATTCAATTAAAATAACCACACATGGCAAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSPRGSSLRGPRPRIPVLVSCQ
PVKGHGTLGESPMPFKRVFCQDGDNVRSCVCAVFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTGTATAAAGACCTAACACTGCTGACC**ATG**ATCAGCGCAG
 CCTGGAGCATCTCCTCATCGGGACTAAAATTGGGCTGTTCTCAAGTAGCACCTCTATCAG
 TTATGGCTAAATCCTGTCCATCTGTGTGCGATGCGGGTTTCATTACTGTAATGATC
 GCTTCTGACATCCATTCCAACAGGAATACCAGAGGATGCTACAACCTCTACCTTCAGAAC
 ACCAAATAATAATGCTGGGATTCTTCAGATTGAAAAACTTGCTGAAAGTAGAAAGAATAT
 ACCTATACCACAAACAGTTAGATGAATTCTACCAACCTCCAAAGTATGTAAGAGTTAC
 ATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAAAAATTCCCTATCTGG
 AAGAATTACATTAGATGACAACCTCTGCAGTTAGCATAGAAGAGGGAGCATTCCGAG
 ACAGCAACTATCTCCGACTGCTTCTGTCCCCTGTAATCACCTAGCACAATTCCCTGGGGTT
 TGCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTCATCACCCT
 CTCTTCAAGGTCTCACTAGTCTAAAACGCCCTGGTTCTAGATGGAAACCTGTTGAACAATCATG
 GTTAGGTGACAAAGTTCTTCAACCTAGTTAATTGACAGAGCTGTCCTGGTGCAGGAATT
 CCCTGACTGCTGCACCAGTAAACCTCCAGGCACAAACCTGAGGAAGCTTATCTCAAGATA
 ACCACATCAATCGGGTCCCCCAAATGCTTTCTTATCTAAGGCAGCTATCGACTGGATA
 TGTCCAATAATAACCTAAGTAATTACCTCAGGGTATCTTGTGATTGGACAATATAACAC
 AACTGATTCTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATGGGTACGTGACTGGT
 TACAATCACTACCTGTGAAGGTCAACGTGCGTGGCTCATGTGCCAAGCCCCAGAAAAGGTTC
 GTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTGATTGTAAGGACAGTGGGATTGTAA
 GCACCATTCAAGATAACCACGTCAATACCAACACAGTGTATCCTGCCAAGGACAGTGGCCAG
 CTCCAGTGACCAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAG
 GGAGTCCCTCAAGAAAAACAATTACAATTACTGTGAAGTCTGTCACCTCTGATACCATTATA
 TCTCTGGAAACTGCTCTACCTATGACTGCTTGGACTCAGCTGGCTAAACTGGCCATA
 GCCCGGCATTGGATCTATAACAGAAACAATTGTAACAGGGAACGCAGTGAGTACTGGTCA
 CAGCCCTGGAGCCTGATTCAACCTATAAAGTATGCATGGTCCATGGAAACCAGCAACCTCT
 ACCTATTGATGAAACTCCTGTTGATTGAGACTGAAACTGCACCCCTCGAATGTACAACC
 CTACAACCACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTACCTTGG
 CTGCCATCATTGGTGGGGCTGTGCCCTGGTTACCATTGCCCTTGTGCTTAGTGTGTTGGT
 ATGTTCATAGGAATGGATCGCTCTCAAGGAACGTGCAATAGCAAAGGGAGGAGAAGAA
 AGGATGACTATGCAGAAGCTGGCACTAAGAAGGACAACCTATCCTGGAAATCAGGGAAACTT
 CTTTCAGATGTTACCAATAAGCAATGAACCCATCTCGAAGGAGGAGTTGTAATACACACCA
 TATTTCCTCTTAATGGAATGAATCTGTACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAA
 GCTACAGAGACAGTGGTATTCCAGACTCAGATCACTCACACTCA**TGA**TGCTGAAGGACTCACA
 GCAGACTTGTGTTGGTTAAACCTAAGGGAGGTGATGGT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCNDRFLTSIPTGIPEDATTLYLQNNQINNAGIPS D LKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKI PYLEELHLD DNSVSAVSIEEGAFRDSNYLRLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSLQGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLP GTNLRKLYLQDNHINRVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPVKVNVRG LMCQAPEKVRGMAIKDLNAELFDCKD S GIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD IKNP KLT KDQQTTGSPSRKTITITVKS VTS DTIHISWKLALPM TALRLSWKLGHSPA FGSITETIVTGERSE YLVT ALEPDSPYKVC M VPMETSNLYLFDETPVC IETETAPLRMYNPTTLNREQEKEPYKNPNLPLAAIIIGGAVALVTIALLALVCWYVHRNGSLFSRN CAYSKGRRRKDDYAEAGTKKD NSILEIRETSFQMLPISNEPI SKEEFVIHTIFPPNGMNLYKNNHSES SSNRSYR DSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561, 640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTCCCAGAGTCCTTGCCAGGCCACCCAGGCTTCTGG
 CAGCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAG
 GCGTGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGA
 GCAGAGGCCGGCATGGCCAGCCTGGGCTGCTCCTGCTCTTACTGACAGCACTGCCACC
 GCTGTGGTCCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCT
 GATCCTGTCTGCGCTGGAGAGAGCCACCGTCTTAGAACAGAGGCTGCCTGAAATCAACCT
 GGATGGCATGGTGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGC
 CCAGGAGCCCTGCTGCAGCCGCTGAGCCTGCGGTGGGATGCTGGGGAGAAGCTGGAGGC
 TGCCATCCAGAGATCCCTCACTACCTCAAGCTGAGTGATCCAAGTACCTAACAGAGAGTTCCA
 GCTGACCCCTCCAGCCCAGGTTTGGAAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTT
 GGTGTACCCCACGTTGGGCCCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGT
 GCAGCTGCTGGGAACCGGGACGGACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAG
 CCTCATGACCAAGCCGGCTGCTCAGGCTACTGCCTGTCCCACCAACTGCTCTTCCCTCTG
 GGCCAGAATGAGGGATGCACACAGGGACCACTCCAACAGAGCCAGGACTATATCACCTCTT
 CTGCGCCAACATGATGGACTTGAACCGCAGAGCTGAGGCCATGGATAACGCCTACCCCTACCG
 GGACATCTTCATGGAAAACATCATGTTCTGTGGAATGGCGGCTCTCCGACTTCTACAAGCT
 CCGGTGGCTGGAGGCCATTCTCAGCTGGCAGAACAGCAGGAAGGATGCTCGGGAGCCTGA
 TGCTGAAGATGAAGAATTATCTAAAGCTATTCAATATCAGCAGCATTTCGAGGAGAGTGAA
 GAGGCAGAAAAACAATTCCAGATTCTCGCTCTGGCTCAGGCTGGAGTACAGTGGCGCAA
 TCTCGGCTCACTGCAACCTTGCCTCTGGTTCAAGCAATTCTTGCCTCATCCTCCGAG
 TAGCTGGACTACAGGAGCGTGCCACCACCTGGCTAATTTTATATTTTTAGTAGAGAC
 AGGTTTCATCATGTTGCTCATGCTGGTCTCGAACCTCTGATCTCAAGAGATCCGCCACCTC
 AGGCTCCAAAGTGTGGATTTAGGTGTGAGCCACCGTGTGGCTGAAAAGCACTTCAAA
 GAGACTGTGTTGAATAAAGGGCAAGGTTCTGCCACCCAGCAGCATTGGGGCTCTCTCCCC
 TAGATGGCTGCTCCTCCCACAACACAGCCACAGCAGTGGCAGCCCTGGGTGGCTCCTATACA
 TCCTGGCAGAATACCCCCCAGCAAACAGAGAGGCCACACCCATCCACACCGCCACCACCAAGCA
 GCCGCTGAGACGGACGGTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCCTTAGTCCTCA
 TCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGGATAAGCAAA
 GCCACCCGACACCAATCTGGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGGGCCGGG
 AGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAACTGAAAAAA

FIGURE 134

MSARGRWEGGRRACRGLGLARAQGAERVTSSEQRPAMASLGLLLLLTLPPPLWSSSLPG
LDTAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSvrekWAQEPLLQPL
SLRVGMLGEKLEAAIQRSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQ
DSFSEERSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQ
GPLQQSQDYINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILS
WQKQQEGCFGEPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLP
PGFKQFSCLILPSSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTC**ATG**GC GGCCGCTCTGTGGGGCTTCTTCCCCTGCTGC
TGCTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGCTGCTGCTGAGGGATCGG
GAGGGAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGCGTGCAGTTGTTCCAGGGG
TGAAGCCTCAGGACTGGATCTCGCGGCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTT
TCCTTAAGACAGATGGGAGTTTGTGGTTCATGATATACTTCTGGATCTTATGTAGTGGAAAG
TTGTATCTCCAGCTTACAGATTGATCCCGTTGAGTGGATATCACTTCGAAAGGAAAATGA
GAGCAAGATATGTGAATTACATCAAAACATCAGAGGTGTCAGACTGCCCTATCCTCTCCAAA
TGAAATCTCAGGTCCACCTTCTTACTTATTAAAAGGAATCGTGGGGCTGGACAGACTTTC
TAATGAACCCAATGGTTATGATGATGGTTCTCCTTATTGATATTGTGCTTCTGCCTAAAG
TGGTCAACACAAGTGATCCTGACATGAGACGGAAATGGAGCAGTCAATGAATATGCTGAATT
CCAACCATGAGTTGCCTGATGTTCTGAGTTCATGACAAGACTCTCTCTCAAAATCATCTG
GCAAATCTAGCAGCGGCAGCAGTAAAACAGGAAAAGTGGGGCTGGCAAAAGGAGG**TAG**TGAG
GCCGTCCAGAGCTGGCATTGACAAACACGGCAACACTGGGTGGCATCCAAGTCTGGAAAAA
CCGTGTGAAGCAACTACTATAAAACTTGAGTCATCCGACGTTGATCTCTTACAACGTGTATGTT
AACTTTTAGCACATGTTGTACTGGTACACGAGAAAACCCAGCTTCATCTTGTCTGT
ATGAGGTCAATATTGATGTCACTGAATTAAATTACAGTGTCTATAGAAAATGCCATTAATAAA
TTATATGAACTACTATACATTATGTATATTAAATTAAAACATCTTAATCCAGAAATCAAAAAAA
AAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSDVQSSEVPGAAAEGSGGSGVGIGDRFKIEGRAVPGVKPQDWIS
AARVLVDGEEHVGFLKTGSFVVHDIPSGSYVVEVSPAYRFDPVRVDITSKGKMRARYVNYI
KTSEVVRLPYPLQMKGSSGPPSYFIKRESWGTDFLMNPMVMMVLPLLIFVLLPKVVNTSDPD
MRREMEQSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGKGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTCTGTGAGATTGATTCGATTCTCCCCAGTTCCCTGTGGTCTGAGGG
 GACCAGAAGGGTGAGCTACGTTGGCTTCTGGAAGGGGAGGCTAT**ATG**CGTCAATTCCCCAAA
 ACAAGTTTGACATTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTT
 CCAGGCCTTACCTGCTGGCACTAACGGCGGAGCCAGGATGGGACAGAATAAAGGAGGCCACG
 ACCTGTGCCACCAACTCGCACTCAGACTCTGAACCTAGACCTGAAATCTTCTCTCACGGGAG
 GCTTGGCAGTTTCTTACTCCTGTGGTCTCCAGATTCAAGGCCTAAGATGAAAGCCTCTAGT
 CTTGCCTTCAGCCTCTCTGCTGCGTTATCTCCTATGGACTCCTCCACTGGACTGAAG
 ACACTCAATTGGGAAGCTGTGTGATGCCACAAACCTTCAGGAAATACGAAATGGATTCT
 GAGATACGGGGCAGTGTGCAAGCCAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACT
 GAGTCTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCCTGCGCCATTGCTAAGACTC
 TATCTGGACAGGGTATTAAAAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGC
 AGCCTGCCAATTCTTCTTACCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACA
 TGCCATTGTGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCACTTGAAAAGCTG
 GAACCTCAGGCAGCAGTTGTGAAGGGTTGGGGAACTAGACATTCTCTGCAATGGATGGAG
 GAGACAGAA**TAG**GAGGAAAGTGATGCTGCTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCT
 TCAATACCTGCAGAGGAGGCATGACCCAAACCACCATCTCTTACTGTACTAGTCTTGTGCT
 GGTACAGTGTATCTTATTGCATTACTTGCTTCCCTGCATGATTGTCTTATGCATCCCC
 AATCTTAATTGAGACCATACTGTATAAGATTGGTAATATCTTCTGCTATTGGATATATT
 TATTAGTTAATATATTATTATTGGCTATTAAATGTATTATTGGACATG
 AAACCTTAAAAAAATTCACAGATTATTTATAACCTGACTAGAGCAGGTGATGTATTGGAT
 ACAGTAAAAAAAAACCTTGTAAATTCTAGAAGAGTGGCTAGGGGGTTATTCAATTGTAT
 TCAACTAAGGACATATTACTCATGCTGATGCTCTGTGAGATATTGAAATTGAAACCAATGAC
 TACTTAGGATGGGTGTGGAATAAGTTGATGTGGAATTGCACATCTACCTACAATTACTG
 ACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTCCAGCCAGGAATCCTACACGG
 CCAGCATGTATTCTACAAATAAGTTCTTGCATACCAAAAAAAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPE
IFSSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQE
IRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHY
TLRKISSLANSFLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDI
LLQWMEETE

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGTCGGTCCGCATCCAG
 CCTAGCGTGTCCACGATGCGGCTGGCTCCGGACTTCGCTACCTGTTGCGTAGCGATCGAG
 GTGCTAGGGATCGCGGTCTCCTCGGGATTCTTCCCAGCCTCGGCTGGAGCCAGTTCTAACTGGACC
 GCGAACACGGAGCAGCCCCAGCGCCCAGCACACTACAGCTGGAGCCAGTTCTAACTGGACC
 ACGCTGCCACCACCTCTTCAGTAAAGTTATTGTTCTGATAGATGCCTTGAGAGATGAT
 TTTGTGTTGGGTCAAAGGGTGTGAAATTATGCCCTACACAACCTACCTGTGGAAAAAGGA
 GCATCTCACAGTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCCTGAATCAAGGCA
 TTGATGACGGGGAGCCTCCTGGCTTGTGACGTACAGGAACCTCAATTCTCCTGCACTG
 CTGGAAGACAGTGTGATAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTATGGAGAT
 GAAACCTGGTTAAATTATTCCCAAAGCATTGTGGAATATGATGGAACAACCTCATTTC
 GTGTCAGATTACACAGAGGTGGATAATAATGTCACGAGGCATTGGATAAAGTATTAAAAAGA
 GGAGATTGGGACATATTAATCCTCCACTACCTGGGCTGGACCACATTGCCACATTCAGGG
 CCCAACAGCCCCCTGATTGGCAGAAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCAC
 ACCTCACTGCAGTCGAAGGAGAGAGACGCCATTACCAATTGCTGGTCTTGTGGTGA
 CATGGCATGTCTGAAACAGGAAGTCACGGGCTCCTCCACCGAGGAGGTGAATACACCTCTG
 ATTTAATCAGTTCTGCTTGAAGGAAACCCGGTGAATACGACATCCAAGCACGTCCAA
TAGACGGATGTGGCTGCGACACTGGCGATAGCACTGGCTTACCGATTCCAAGACAGTGT
 GGGAGCCTCCTATTCCAGTTGTGGAAGGAAGACCAATGAGAGAGCAGTTGAGATTTC
 TTGAATAACAGTGCAGCTTAGTAAACTGTTGCAAGAGAATGTGCCGTATGAAAAAGATCCT
 GGGTTGAGCAGTTAAAATGTCAGAAAGATTGCATGGAACTGGATCAGACTGTACTTGGAG
 GAAAAGCATTCAAGTCCTATTCAACCTGGGCTCCAAGGTTCTCAGGCAGTACCTGGATGCT
 CTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAGTTCTCACCTGCTCCTGCTCA
 GCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCAGTGCATCTCCTGGTTTT
 CTCTGCTCTTATTGGTGAATCCTGGTCTTTCGGCGTTACGTCAATTGTGTGCACCTCAG
 CTGAAAGTTCGTGTACTTCTGTGGCTCTCGTGGCTGGCAGGCTGCCTTC
 GACTCTGGTTGAACACCTGGTGTGCCAGTGCTGGCAGTGCCTGGACAGGGGCCTCAGG
 GAAGGACGTGGAGCAGCCTATCCCAGGCCTCTGGGTGTCCGACACAGGTGTTCACATCTGT
 GCTGTCAGGTCAAGTGCCTCAGTTGGAAAGCTAGGTTCTGCGACTGTTACCAAGGTGAT
 TGTAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGAATCGGACA
 GCCTCCCAGCAGAGGTGTGGAGCTGCAGTGAGGAAGAGACAATCGGCCTGGACACTC
 AGGAGGGTCAAAAGGAGACTTGGTCGACCACTCATCCTGCCACCCCCAGAATGCATCCTGCC
 TCATCAGGTCCAGATTCTCAAGGCGGACGTTCTGTGGAATTCTTAGTCCTGGCCT
 CGGACACCTCATTGTTAGCTGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCA
 CACTCAGATCCACAGAGGCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCC
 CCCACCCCAACCCTGCACAGCCCTCATCCCCTCTGGCTTGAGGCCAGAGGCCCTGTGCTG
 AGTGTCTGACCGAGACACTCACAGCTTGTACAGGGCACAGGCTTCCTCGGAGCCAGGATG
 ATCTGTGCCACGCTTGCACCTCGGGCCATCTGGCTCATGCTCTCCTGCTATTGAATT
 AGTACCTAGCTGCACACAGTATGTAGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTLPPP
LFSKVIVLIDALRDDVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGS
LPGFVDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTSFFVSDYT
EVDNNVTRHLDKVLKRGDWLILHHLGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQS
KERETPLPNLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTATCGTGACGCACCTGAAAGTCTGAGAGCTACTGCCCTA
CAGAAAGTTACTAGTGCCTAAAGCTGGCGCTGGCACTG**ATG**TACTGCTGCTGGAGTAC
AACTTCCCTATAGAAAACAACGCCAGCACCTTAAGACCACACACCTCAGAGTGAAGAAC
TTAAACCCGAAGAAATTCAAGCATTGACCAAGGATCACAAAGTACTGGTCCTGGACTCTGGG
AATCTCATAGCAGTCCAGATAAAACTACATAGCCAGAGATCTTCTTGCATTAGCCTCA
TCCTTGAGCTCAGCCTCTGGAGAAAGGAAGTCCGATTCTCCTGGGGTCTCTAAAGGGAG
TTTGTCTCTACTGTGACAAGGATAAAGGACAAAGTCATCCATCCCTCAGCTGAAGAAGGAG
AAACTGATGAAGCTGGCTGCCAAAAGGAATCAGCACGCCGCCCTCATCTTATAGGGCT
CAGGTGGGCTCCTGGAACATGCTGGAGTCGGCGCTCACCCGGATGGTCATCTGCACCTCC
TGCAATTGTAATGAGCCTGTTGGGGTGACAGATAAATTGAGAACAGGAAACACATTGAATT
TCATTTCAACCAGTTGCAAAGCTGAAATGAGCCCCAGTGAGGTAGCGAT**TAG**GAAACTGCC
CCATTGAACGCCCTCGCTAATTGAACTAATTGTATAAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLEYNFPIENNQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLDGNLIAVPDKNYIRP
EIFFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESAR
RPFIFYRAQVGWSWNMLESAAHPGWFICTSCNCNEPVGVTDFENRKHIEFSFQPVCKAEMS
EVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.

amino acids 50-55, 87-92

Interleukin-1

amino acids 37-182

FIGURE 143

CTAGAGAGTATAAGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTT
AATCCAGGATCCTGTCCCTGTCTGTAGGAGTGCCTGCCAGTGTGGGTGAGACAAG
TTTGTCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACTCC
TGTGGGCACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACC**ATG**CTGGGCC
TGCCGTGGAAGGGAGGTCTGTCCTGGCGCTGCTGCTGCTCTCTTAGGCTCCAGATCCTGC
TGATCTATGCCTGGCATTCCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTT
ACCTCCCTGCCACAGTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATG
CCTACAGACTGGGCACATCTTGAATTCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCT
CAATGGAGCTACTGCTGGGAGAACTAGGTGTGGAAATTGAAGACGACATTGACAACGTGCC
ATTTCCAAGAAAGCACAGAGCTGAACAAACTTACCTGCTTCTCACCACAGCACCAGGC
CCTGGATGACTCAGTCAGCCTCCTGAACAAAGACCTGCTGGAGGGATTCCAC**TGA**GTGAAAC
CCACTCACAGGCTTGTCCATGTGCTGCTCCCACATTCCGTGGACATCAGCACTACTCTCCTGA
GGACTCTCAGTGGCTGAGCAGCTTGGACTTGTGTTATCCTATTTGCATGTGTTGAGA
TCTCAGATCAGTGTGTTAGAAAATCCACACATCTGAGCCTAACATGTAGTGTAGATCATTA
AACATCAGCATTAAAGAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA

FIGURE 144

MLGLPWKGGLSWALLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQOSK
DYYAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNNTFTCFFTI
STRPWMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGGCCAAGGTTCTGACGCGATGAGGA
AGCACCTGAGCTGGTGGCTGGCCACTGTCTGCATGCTGCTCTCAGCCACCTCTGC
TCCAGACGAGGGGCATCAAGCACAGAAATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCC
AGATCACTGAGGCCAGGTGGCTGAGAACCGCCGGAGCCTCATCAAGCAAGGCCGCAAGC
TCGACATTGACTTCGGAGCCGAGGGCACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCG
ATGGCATCCACTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTCACC
GCATCAATGCCACCCAGGCGCGAACCAAGGGGAGTCCAGAACGCCAGACAACAAGCTCCACC
AGCAGGTGCTCTGGCGGCTGGTCCAGGAGCTTGCTCCCTCAAGCATTGCGAGTTGGTTGG
AGAGGGCGCAGGACTTCGGTCACCATGCACCAGCCAGTGCTCCTGCCTCTGGCTTGA
TCTGGCTCATGGTAAATAAGCTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAA
TCCTGGCAAGTGACCCAGCTCTCCCCAAACCCACGCGTGTGAAGGTGCCAGGAGC
GGCGATGCACTCGCACTGCAAATGCCGCTCCCACGTATGCCCTGGTATGCGCTGC
GATAGATGGGGACTGTGGCTTCTCCGTCACTCCATTCTCAGCCCTAGCAGAGCGTCTGGCA
CACTAGATTAGTAGTAAATGCTTGATGAGAACACATCAGGCAGTGCGCCACCTGCTTCAC
AGTACTTCCAACAACCTTAGAGGTAGGTATTCCGTTACAGATAAGGAAACTGAGGC
CCAGAGAGCTGAAGTACTGCACCCAGCATCACCAGCTAGAAAGTGGCAGAGCCAGGATTCAAC
CCTGGCTTGTCTAACCCAGGTTCTGCTCTGTCCAATTCCAGAGCTGTGGTATC
TATGTCTCACAGGGACCCACATCCAAACATGTATCTCTAATGAAATTGTGAAAGCTCCATGTT
TAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQG
RKLDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNK
LHQQVLWRLVQELCSLKHCEFWLERGAGLRVTMHQPVLLCLLALIWLVMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGCGTGACCACCATGTCTGGTCCAGAGTCTCAT
TTCCTGATGATTATAGACTCAAAGAAAACTCA**TGTT**CAGAAGCTCTCTCTGGCCTC
CTCTCTGTCTTCTTCCCTTTCTTCTTATTAAATTAGTAGCATCTACTCAGAGTCATGCA
AGCTGGAAATCTTCATTTGCTTGTCAAGTGGGTAGGTCACTGAGTCTAGTTTATT
TGAAATTCAACTTCAGATTCAAGGGTACATGTGAAGGTTGTTATGAGTATATTGCA**TG**
GATGCTGAGGTTGGGT

148/168

FIGURE 148

MFRSSLLFWPPLCLLSLFLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGT
CEGLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTTACCTGGAGAC
 TTGACTCCCGCGCCCCAACCTGCTTATCCCTGACCGTCGAGTGTAGAGATCCTGCAGC
 CGCCCAGTCCCGGCCCTCTCCGCCAACACCCACCCTGGCTCTCCTGTTACTCC
 TCCTTTCATCATAACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGCTGTGACCCAAGCC
 GAGCGTGGAAAGAATGGGGTTCTCGGGACCGGCATTGGATTCTGGTTAGTGCTCCGATT
 CAAGCTTCCCCAACCTGGAGGAAGCCAAGACAAATCTACATAATAGAGAATTAAGTGCA
 GAAAGACCTTGAATGAACAGATTGCTGAAGCAGAAGACAAGATTAAAAAACATATCCT
 CCAGAAAACAAGCCAGGTAGAGCAACTATTCTTGATAACTGAACTGCTAAAGGCA
 ATAACAGAAAAGGAAAAATTGAGAAAGACAATCTATAAGAAGCTCCCCACTGATAAT
 AAGTTGAATGTGGAAAGATGTTGATTCAACCAAGAATCGAAAAGTGATCGATGATTATGACTCT
 ACTAAGAGTGGATTGGATCATAAATTCAAGATGATCCAGATGGTCTCATCAACTAGACGGG
 ACTCCTTAACCGCTGAAGACATTGCCATAAAATCGCTGCCAGGATTATGAAGAAAATGAC
 AGAGCCGTGTTGACAAGATTGTTCTAAACTACTTAATCTGGCCTTATCACAGAAAGCCAA
 GCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAATTAATCTCAAAGGAAGCCAAC
 AATTATGAGGAGGATCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAATACCA
 GAGAAAGTGACTCCAATGGCAGCAATTCAAGATGGCTTGCTAAGGGAGAAAACGATGAAACA
 GTATCTAACACATTAAACCTTGACAAATGGCTTGAAAGGAGAAACTAAAACCTACAGTGAAGAC
 AACTTGAGGAACCTCAATATTCCAAATTCTATGCGCTACTGAAAAGTATTGATTGAGAA
 AAAGAACAAAAGAGAAAGAACACTGATTACTATCATGAAAACACTGATTGACTTGTGAAG
 ATGATGGTGAATATGGAACAATATCTCCAGAAGAAGGTGTTCTACCTTGAAAACCTGGAT
 GAAATGATTGCTCTCAGACCAAAACAAGCTAGAAAAATGCTACTGACAATATAAGCAAG
 CTTTCCCAGCACCATTAGAGAAGACTCATGAAGAACAGACAGTACCAAGGAAGCAGCT
 AAGATGGAAAAGGAATATGGAAGCTTGAGGATTCCACAAAGATGATAACTCCAACCCAGGA
 GGAAAGACAGATGAACCCAAAGGAAAACAGAACGCCTATTGGAAGCCATCAGAAAAATATT
 GAATGGTTGAAGAACATGACAAAAAGGGAAATAAGAAGATTATGACCTTCAAAGATGAGA
 GACTTCATCAATAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTGACAAGGAAGAAGCC
 GAGGCCATCAAGCGCATTAGCAGCCTGTAAAATGGCAAAAGATCCAGGAGTCTTCAAC
 TGTTTCAGAAAACATAATATAGCTAAAACACTTCTAATTCTGTGATTAAAATTTTGACCC
 AAGGGTTATTAGAAAGTGCTGAATTACAGTAGTTAACCTTACAAGTGGTTAAAACATAGC
 TTTCTTCCGTAAAACATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAA
 AAA

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENK
PGQSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSG
LDHKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTL
EDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPFKVTPMAAIQDGLAKGENDETVSNT
LTLTNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFKMMVK
YGTISPEEGVSYLENLDEMIALQTKNKLEKNATDNISKLFPAPSEKSHEETDSTKEEAAKMEK
EYGSLKDKSTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGKEDYDLGMRDFIN
KQADAYVEKGILDKEEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTC
AAGATGGTCCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTAT
 CTGCATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAAGGTCACTAAAGGTGAAGAG
 ATCAGCGTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCGTCATCCTGGGTGTCCAG
 GGTGGAAGCCAGTGCCTGTCATGTGGGTGGGCAGGAGCCGACTCTAACACTAGAGCCAGTG
 AACATCATGGAGCTCTATCTTGGTGCAGGAATCCAAGAGCTTCACCTCTACCGGCGGGAC
 ATGGGGCTCACCTCCAGCTCGAGTCGGCTGCCTACCCGGCTGGTCTGTGCACGGTGCCT
 GAAGCCGATCAGCCTGTCACTCACCCAGCTTCCCAGAATGGTGGCTGGAATGCCCCCATC
 ACAGACTTCTACTTCCAGCAGTGTGACTAGGGCAACGTGCCCCCCAGAACTCCCTGGCAGAG
 CCAGCTCGGGTGAGGGGTGAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCTCAG
 GACCCCCACGTCTGACTTAGTGGCACCTGACCACTTGTCTCTGGTCCAGTTGGATAA
 ATTCTGAGATTGGAGCTCAGTCCACGGTCCCTCCCCACTGGATGGTCTACTGCTGTGGAAC
 CTTGTAAAAACCATGTGGGGTAAACTGGGAATAACATGAAAAGATTCTGTGGGGGTGGGTG
 GGGGAGTGGTGGGAATCATTCTGCTTAATGGTAACTGACAAGTGTACCCCTGAGCCCCCAG
 GCCAACCCATCCCCAGTTGAGCCTTATAGGGTCAGTAGCTCCACATGAAGTCCTGTCACTC
 ACCACTGTGCAGGAGAGGGAGGTGGTCAAGAGTCAGGGATCTATGCCCTGGCCCAGCCCC
 ACCCCCCTCCCTTAATCCTGCCACTGTCAATGCTACCTTCTATCTCTCCCTCATCATC
 TTGTTGTGGCATGAGGAGGGTGGTGTAGTCAGAAGAAATGGCTCGAGCTCAGAAGATAAAAGA
 TAAGTAGGGTATGCTGATCCTCTTTAAAAACCCAAGATAACATCAAATCCAGATGCTGGT
 CTCTATTCCCATGAAAAAGTGTCACTGACATATTGAGAAGACCTACTTACAAAGTGGCATATA
 TTGCAATTATTAAATTAAAAGATACCTATTATATTTCTTATAGAAAAAAAGTCTGGAA
 GAGTTTACTCAATTGTAGCAATGTCAGGGTGGCAGTATAGGTGATTTCTTTAATT
 TGTTAATTATCTGTATTCCTAATTTCTACAATGAAGATGAATTCTTGTATAAAATAA
 GAAAAGAAATTAAATCTTGAGGTAAGCAGAGCAGACATCATCTGATTGTCTCAGCCTCCAC
 TTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGGTAGTAGTGATCA
 GGAAACAGATCTCAGCAAAGCCACTGAGGAGGGCTGTGCTGAGTTGTGGCTGGAATCT
 CTGGGTAAAGGAACTTAAAGAACAAAAATCATCTGGTAATTCTTCTAGAAGGATCACAGCCC
 CCTCTCAAATTCACATCCTCTTGAATCTCAGTCTGTGAGTTATTGGAGATAAGGTCTCTG
 CAGATGTAGTTAGTTAAGACAAGGTCACTGGATGAAGGTAGACCTAAATTCAATATGACTG
 GTTCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCAGGGAAAGACTATGAAAG
 ATGAAGGCAGAGATCGGAGTTGCAAGCCACAAGCTAAGAACACCAAGGATTGTGGCAACCA
 TCAGAAGCTTGAAGAGGCAAAGAAGAATTCTCCCTAGAGGCTTAGAGGGATAACGGCTCT
 GCTGAAACCTTAATCTCAGACTCCAGCCTCTGAACGAAGAAAGAATAAATTCTGGCTGTT
 TAAGCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAAACTAATACAGCTGCTAAATGATC
 CCTGTCTCCTCGTGTTCATTCTGTGTGTCCTCCCACAATGTACCAAAGTTGTCTTGT
 TGACCAATAGAATATGGCAGAAGTGTGATGGCATGCCACTTCCAAGATTAGGTATAAAAGACAC
 TGCAGCTCTACTTGAGGCCCTCTCTGCCACCCACCGCCCCAATCTATCTTGGCTCACT
 CGCTCTGGGGAGCTAGCTGCCATGCTATGAGCAGGCCTATAAGAGACTTACGTGGTAAA
 AATGAAGTCTCCTGCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATC
 GATGTTGTTGTTAAGTTGCTCAGTTGGCTAACCTGTTATGCAGCAATAGATAAATAA
 TATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEEISVVPNRWLDASLSPVILGVQG
GSQCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPE
ADQPVRTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCA**ATGGCC**
GCCCTGCAGAAATCTGTGAGCTTTCTTATGGGGACCCTGGCCACCAGCTGCCTCCTCTC
TTGGCCCTCTGGTACAGGGAGGAGCAGCTGCCCATCAGCTCCCCTGCAGGCTTGACAAG
TCCAACCTCCAGCAGCCCTATATCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTTG
GCTGATAACAACACAGACGTTCGTCTCATGGGGAGAAACTGTTCCACGGAGTCAGTATGAGT
GAGCGCTGCTATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAGAAGTGCTGTTCCCTCAA
TCTGATAGGTTCCAGCCTTATATGCAGGAGGTGGTGCCTCCTGCCAGGCTCAGCAACAGG
CTAACGCACATGTCATATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGAAAAGCTGAAG
GACACAGTAAAAAGCTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTG
TTTATGTCTCTGAGAAATGCCTGCATT**TGA**CCAGAGCAAAGCTGAAAAATGAATAACTAACCC
CCTTCCCTGCTAGAAATAACAATTAGATGCCCAAAGCGATTTTTAACCAAAAGGAAGA
TGGGAAGCCAAACTCCATCATGATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGG
AAACCAATGCCACTTTGTTATAAGACCAGAAGGTAGACTTCTAACGATAGATATTATTG
ATAACATTTCATTGTAAGTGGTGTCTATACACAGAAAACAATTATTAAATAATTGTC
TTTTCCATAAAAAAGATTACTTCCATTCTTAGGGAAAAAACCCCTAAATAGCTTCATG
TTCCATAATCAGTACTTATATTATAATGTATTATTATTATAAGACTGCATT
TTATATCATTATTATAATATGGATTATTATAAGAACATCATTGCTACTTGAGTG
TAAGGCTAATATTGATATTGACAATAATTATAGAGCTATAACATGTTATTGACCTCAA
TAAACACTGGATATCCC

FIGURE 154

MAALQKS VSSFLMGT LAT SCLLL ALLVQGGAA APISSH CRLDKS NFQQPY ITNRTF MLAKEA
SLADNN TDVRLIGEKL FHGVSM SERCYLMK QVLNFT LEEVLF PQSDRF QPYM QEVVPFLARLS
NRLSTCHIEGDDLH IQRNVQKL KDTVKKLGE SGEIKAIGEL DLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTCCACGAGGCCT
GTCAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCCAGCATGTACCAGGTAGTCAGTCAGA
GGGCTGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGG
CCAAGCTGCCAGGTTGGGCTGGGCCAAGTGGAGTGAGAAACTGGGATCCCAGGGGGAGG
GTGCAGATGAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCATTAGCCTTCCTACAG
GTGGTTGCATTCTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGC
CCCAGCAAAGGGCAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCAGTGCCTGTGCCTCCC
CTAGAGCCTGCTAGGCCAACGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTC
AACAGCAGGGCCATCTCCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCAG
GACCTGTACCACGCCCGTGCCTGTGCCGCACTGCGTCAGCCTACAGACAGGCTCCCACATG
GACCCCCGGGCAACTCGGAGCTGCTCTACCACAACCAGACTGTCTTACAGGCGGCCATGC
CATGGCGAGAAGGGACCCACAAGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTA
GCTTGTGTGTGTGCAGGCCGTGTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTT
TTGGGAAACCTGGAGCCAGGTGTACAACCACCTGCCATGAAGGGCCAGGATGCCAGATGCTT
GGCCCGTGTGAAGTGCTGTCTGGAGCAGCAGGATCCCAGGACAGGATGGGGGCTTGGGAA
AACCTGCACCTCTGCACATTGAAAAGAGCAGCTGCTGCTTAGGGCCCGGAAGCTGGTGT
CCTGTCATTTCTCTCAGGAAAGGTTCAAAGTTCTGCCATTTCTGGAGGCCACCAACTCCT
GTCTCTCCTCTTCCCACCCCTGCTACCCCTGGCCCAGCACAGGCACTTCTAGATATTTC
CCCCTGCTGGAGAAGAAAGAGGCCCTGGTTATTGTTACTCATCACTCAGTGAGC
ATCTACTTGGGTGCATTCTAGTGTAGTTACTAGTCTTGTACATGGATGATTCTGAGGAGGA
AGCTGTTATTGAATGTATAGAGATTATCCAAATAAATATCTTATTAAAAATGAAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGTHTYSHWPSCCPKGQDTSEELLRWSTVPVPPL
PARPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDP
RGNSELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCG**ATG**TCGCTCGTGTGCTAAGCCTGGCCGCGCTGTGCAGGAGGCCGTACCCCGAGAG
 CCGACC GTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTA
 ATCCCCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGGAC
 TATTCAATTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGAAG
 GCCACCAAGATTGTGTGACGGGAAAAGCAACTTCCAGTCCTACAGCTGTGAGGTGCAAT
 TACACAGAGGCCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTCTACATC
 GGCTTCCCTGTAGAGCTGAACACAGTCTATTTCATTGGGCCATAATATTCTAATGCAAAT
 ATGAATGAAGATGGCCCTTCCATGTCTGTGAATTCACCTCACAGGCTGCCTAGACCACATA
 ATGAAATATAAAAAAAAGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGT
 AAGAAGAATGAGGAGACAGTAGAACGTACAACCCTCCCCTGGAAACAGATAACATG
 GCTCTTATCCAACACAGCACTATCATGGGTTTCTCAGGTGTTGAGCCACACCAGAAGAAA
 CAAACCGAGCTTCAGTGGTATTCCAGTGACTGGGATAGTGAAGGTGCTACGGTGCAGCTG
 ACTCCATATTCCTACTTGTGGCAGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGC
 CCACAAACAGGCCTTCCCTCTGGATAACAACAAAAGCAAGCCGGAGGCTGGCTGCCT
 CTCCTCCTGCTGTCTGCTGGTGGCACATGGTGCTGGCAGGGATCTATCTAATGTGG
 AGGCACGAAAGGATCAAGAAGACTTCCTTTCTACCACCAACTACTGCCCCCCATTAAGGTT
 CTGTGGTTACCCATCTGAAATATGTTCCATCACACAATTGTTACTTCACTGAATTCTT
 CAAAACCATTGCAGAAGTGAGGTATCCTTGAAAAGTGGCAGAAAAAGAAAATAGCAGAGATG
 GGTCCAGTGCAGTGGCTGCCACTCAAAAGAAGGCAGCAGACAAAGTCGTCTCCTTCTTCC
 AATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCAGTGAGAAC
 TCTCAAGACCTCTTCCCCCTGCCTTAACCTTTCTGCAGTGATCTAAGAACGCCAGATTCT
 CTGCACAAATACGTGGTGGTCTACTTAGAGAGATTGATAAAAAGACGATTACAATGCTCTC
 AGTGTCTGCCCAAGTACCACCTCATGAAGGATGCCACTGCTTCTGTGCAGAACTTCTCCAT
 GTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACGATGGCTGCTCCTTG
TAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTTSVATGDYS
ILMNVSWVLRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGF
PVELNTVYFIGAHNIPNANMNEGPSMSVNFTSPGCLDHIMKYKKCVKAGSLWDPNITACKK
NEETVEVNFTTPLGNRYMALIQHSTIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTP
YFPTCGSDCIRHKGTVVLCPTGVPFPLDNNKSKPGGWLPLLLSLLVATWVLVAGIYLMWRH
ERIKKTSFSTTLLPPIKVLVVYPSEICFHHTICYFTEFLQNHCRSEVILEKWQKKKIAEMGP
VQWLATQKKAADKVVFLLSDVNSVCDGTCGKSEGSPSENSQDLFPLAFNLFCSDLRSQIHLH
KYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVVKQQVSAGKRSQACHDGCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201
and 283 - 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCCCTGCATGCCAGCCATGGTCAAGTACTTGCTG
CTGTCGATATTGGGGCTTGCCTTCTGAGTGAGGCAGCTCGAATCCCCAAAGTAGGA
CATACTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGAC
ATTGGCATCATCAATGAAAACCAGCGCGTTCATGTCACGTAACATCGAGAGCCGCTCCACC
TCCCCCTGGAATTACACTGTCACGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCC
CAGTGTAGGAACCTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTT
CCCATCCAGCAAGAGACCCCTGGTCCGGAGGAAGCACCAAGGCTGCTGTGTTCTTCCAG
TTGGAGAAGGTGCTGGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCACTGAGAAG
TAAGAGGTGCATATCCACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPAMVKYLLSILGLAFLSEAAARKIPKGHTFFQKPESCPPVPGGSMKLDIGIIN
ENQRVSMSRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQE
TLVVRRKHQGCSVSFQLEKVLVTVGCTCVTPVIHHQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGTCAGGACTCCCA
 GGACAGAGAGTGCACAAACTACCCAGCACAGCCCCCTCCGCCCTCTGGAGGCTGAAGAGGG
 ATTCCAGCCCCTGCCACCCACAGACACGGGCTGACTGGGTGTCTGCCCTTGGGGGGGGGG
 CAGCACAGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCCTGTGCCCTGGTTCTT
 GCTGTCTTGGCACTGGCGAAGCCCAGTGGCCTTCTCTGGAGAGGGCTGTGGGCCTCA
 GGACGCTACCCACTGCTCTCCGGGCCTCTCCTGCCGCCTGGGACAGTGACATACTGCCT
 GCCTGGGACATCGTGCCTGCTCCGGGCCGTGCTGGGCCTACGCACCTGCAGACAGAGCT
 GGTGCTGAGGTGCCAGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGC
 CGTGCATGGGCACTGGGAAGAGCCTGAAGATGAGGAAAAGTTGGAGGAGCAGCTGACTCAGG
 GGTGGAGGAGCCTAGGAATGCCTCTCCAGGCCAAGTCGTGCTCTCCTCCAGGCCTACCC
 TACTGCCGCTGCGTCTGCTGGAGGTGCAAGTGCCTGCTGCCCTGTGCAGTTGGTCAGTC
 TGTGGCTCTGGTATATGACTGCTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTC
 CTATACTCAGCCCAGGTACGAGAACGAACTCAACCACACACAGCAGCTGCCCTGCCCTG
 GCTCAACGTGTCAGCAGATGGTGACAACGTGATCTGGTTCTGAATGTCTTGAGGAGCAGCA
 CTTCGGCCTCTCCCTGTAUTGGAAATCAGGTCCAGGGCCCCAAAACCCCGGTGGCACAAAAAA
 CCTGACTGGACCGCAGATCATTACCTGAACCACACAGACCTGGTCCCTGCCTCTGTATTCA
 GGTGTGGCCTCTGGAACCTGACTCCGTTAGGACGAACATCTGCCCTCAGGGAGGACCCCCG
 CGCACACCAGAACCTCTGGCAAGCCGCCGACTGCGACTGCTGACCCCTGCAGAGCTGGCTGCT
 GGACGCACCGTGCTCGCTGCCGCAGAACGGCACTGTGCTGGCGGCTCCGGTGGGACCC
 CTGCCAGCCACTGGTCCCACCGCTTCTGGAGAACGTCACTGTGGACAAGGTTCTGAGTT
 CCCATTGCTGAAAGGCCACCCCTAACCTCTGTGTTAGGTGAAACAGCTCGGAGAACGCTGAGCT
 GCAGGAGTGCTGTGGCTGACTCCCTGGGCCCTCTCAAAGACGATGTGCTACTGTTGGAGAC
 ACGAGGCCCCCAGGACAACAGATCCCTCTGTGCCTGGAACCCAGTGGCTGTACTCACTACC
 CAGCAAAGCCTCCACGAGGGCAGCTCGCCTGGAGAGTACTTACTACAAGACCTGCAGTCAGG
 CCAGTGTGTCAGCTATGGGACGATGACTTGGAGCGCTATGGGCCCTGCCCTGGACAAATA
 CATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTGCCGCTGCGCTTCCCT
 CATCCTCTTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAAACAGGACGTCCG
 CTCGGGGCGGCCAGGGCCGCGCGCTCTGCTGCCCTGTGCCAGCTGCCGCTGCGCGTGGCGTAGA
 CGAGCGCTGGTGGCGCCCTGGCGTGGCCCTGTGCGCTGCCAGCTGCCGCTGCGCGTGGCGTAGA
 CCTGTGGAGCCGTCGTGAACTGAGCGCGCAGGGGCCGTGGCTTACGCGCAGCGGGCG
 CCAGACCCCTGCAGGAGGGCGCGTGGTGGCTTGCTCTCTCCCGTGCCTGGCGCTGTG
 CAGCGAGTGGCTACAGGATGGGTGTCCGGGCCGGCGCACGGCCCGCACGACGCCTCCG
 CGCCTCGCTCAGCTCGTGTGCCGACTTCTGCAGGGCCGGCGCCGGCAGCTACGTGGG
 GCCCTGCTCGACAGGCTGCCACCCGGACGCCGTACCCGCCCTTCCGCACCGTGCCGT
 CTCACACTGCCCTCCAACTGCCAGACTTCTGGGCCCTGCAGCAGCCTCGCGCCCCGCG
 TTCCGGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCGGCCCTTCAGCCAGCCCTGGATAG
 CTACTTCCATCCCCGGGACTCCCGGCCGGACGCCGGTGGGACCAAGGGCGGGACCTGG
 GGCGGGGACGGACTTAAATAAAGGCAGACGCTTTCTAAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSLERLVPQDATHCSPGLSRLWDSDILCLPGDIVPAPGPVLAP
THLQTELVLRCQKETDCDLCLRVAVHLAVHGHWEEPEDEEKFGGAADSGVEPRNASLQAQVV
LSFQAYPTARCVLLEVQVPAALVQFGQSVGSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQ
QLPALPWLNVSADGDNVHLVNVSEEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDL
VPCLCIQVWPLEPDSVRTNICPFREDPRAHQNLWQAARLRLTLQSWLLDAPCSLPAEAALCW
RAPGGDPCQPLVPPLSWENVTVDKVLEFPPLLKGHPNLCVQNSSEKLQLQECLWADSLGPLKD
DVLLLETRGQPQDNRSICALEPSGCTSLPSKASTRAARLGEYLLQDLQSGQCLQLWDDDLGALW
ACPMDKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARGRAALLLY
SADDGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAWFHAQRRQTLQEGGVVLLFS
PGAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPA
LFRTVPVFTLPSQLPDFL GALQQPRAPRSGRLQERAEQVSRALQPALDSYFHPPGT PAPGRGV
GPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCG**ATG**AGGACGCTGCTGACCATCTGACTGTGGGATCCCTGG
 CTGCTCACGCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCCAGTCAGCAACT
 TTGAAAACATCCTGACGTGGGACAGCAGGCCAGAGGGCACCCAGACACGGTCTACAGCATCG
 AGTATAAGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGGATCACCGGA
 AGTCCTGCAACCTGACGGTGGAGACGGCAACCTCACGGAGCTACTATGCCAGGGCACCCTG
 GTCAGTGCAGGCCAGGCCGGTCAGCCACCAAGATGACTGACAGGTTCAGCTCTGCAGCACACT
 ACCCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGAGATCGATTGAGATGATTGTTCAT
 CCTACCCCCACGCCAATCCGTGCAGGCATGGCCACCGGCTAACCTGGAAAGACATCTCCAT
 GACCTGTTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTGGAGGGAAAG
 CAGAGAGAATATGAGTTCTCGGCCTGACCCCTGACACAGAGTTCTGGCACCATCATGATT
 TGCCTCCCACCTGGCCAAGGAGAGTGGCCCTACATGTGCCAGTGAAGACACTGCCAGAC
 CGGACATGGACCTACTCCTCTCCGGAGCCTCTGTTCTCCATGGGCTTCGTGCGAGTA
 CTCTGCTACCTGAGCTACAGATATGTCACCAAGCCGCCTGCACCTCCAACTCCCTGAACGTC
 CAGCGAGTCCTGACTTCCAGCCGCTGCGCTTCATCCAGGAGCACGTCCTGATCCCTGTCTT
 GACCTCAGCGCCCCAGCAGTCTGGCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGA
 CCCAGGGAGCCCGCAGGAGCTCACAGCGGCATAGCCTGTCCAGATCACCTACTTAGGGCAG
 CCAGACATCTCCATCCTCCAGCCCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCC
 TATGCCCAAACGCTGCCCTGAGGTGGGCCCCATCCTATGCACCTCAGGTGACCCCCGAA
 GCTCAATTCCCATTCTACGCCAACAGGCCATCTCTAAGGTCCAGCCTCCTATGCCCT
 CAAGCCACTCCGGACAGCTGGCCTCCCTCCTATGGGTATGCATGGAAGGTTCTGGCAAAGAC
 TCCCCCAGCTGGACACTTCTAGTCCTAAACACCTTAGGCCTAAAGGTCAAGCTTCAGAAAGAG
 CCACCACTGGAAAGCTGCATGTTAGGTGGCCTTCTGCAGGAGGTGACCTCCTGGCTATG
 GAGGAATCCAAGAACAAAATCATTGCACCAAGCCCTGGGATTTGCACAGACAGAACATCT
 GACCCAAATGTGCTACACAGTGGGAGGAAGGGACACCACAGTACCTAAAGGGCCAGCTCCC
 CTCCTCTCCTCAGTCAGATCGAGGGCCACCCATGTCCTCCCTTGCAACCTCCTCCGGT
 CCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGCTGCTGGAGTCCCTGTGTGCCC
 AAGGATGAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTGGAGCAGCCCACAGAACTGGAT
 TCTCTTTCAGAGGCCTGGCCCTGACTGTGCAGTGGGAGTCCT**TGA**GGGAATGGGAAAGGCTT
 GGTGCTTCCTCCCTGTCCTACCCAGTGTACATCCTGGCTGTCAATCCCAGCCTGCCAT
 GCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTGAGAGAAGCAGAGGGAGTGGCATG
 CAGGGCCCTGCCATGGTGCCTCCTCACCGAACAAAGCAGCATGATAAGGACTGCAGCGG
 GGGAGCTCTGGGAGCAGCTGTGTAGACAAGCGCGTGCTCGCTGAGCCCTGCAAGGCAGAAA
 TGACAGTGCAAGGAGGAAATGCAGGGAAACTCCGAGGTCCAGAGCCCCACCTCTAACACCA
 TGGATTCAAAGTGCTCAGGAATTGCCTCTCCTGGCCCTTCCTGGCCAGTTCAAAATCT
 AGCTCGACAGAGCATGAGGCCCTGCCTCTGTCAATTGTTCAAAGGTGGAAAGAGAGCCTG
 GAAAAGAACCAAGGCCAGGGAAAGAACCAAGAGGAGGCTGGGAGAACACCAGAACACTGCACT
 TCTGCCAAGGCCAGGGCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGCTCAT
 TCCCAGCCAGGGCAACTGCCTGACGTTGCACGATTCTGCTCATTGTCATAGAACAAAG
 CGAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTCTGCAGGCAGGAGTTCAGAC
 CCTATCCTGAGAATGGGTTGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGTACAATAA
 CACACTGTACTGATGTCACAACCTTGCAAGCTCTGCCTGGGTTCAGCCCATCTGGCTCAA
 TTCCAGCCTCACCACCTACAAGCTGTGTGACTCAAACAAATGAAATCAGTGCCAGAACCTC
 GGTTCTCATCTGTAATGTGGGATCATAACACCTACCTCATGGAGTTGGTGAAGATGAA
 ATGAAGTCATGTCTTAAAGTGCTTAATAGTGCCTGGTACATGGCAGTGCCAATAACGGT
 AGCTATTTAAAAAAAAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGER
DWVAKKGCQRITRKSCNLTVETGNLTELYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDV
TCISKVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKOREYEFF
GLTPDTEFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFGAFLFSMGFLAVLCYLSYR
YVTKPPAPPNSLNQRVLTQPLRFIQEHVLIPVFDLSGPSSLAQPVQYSQIRVSGPREPAGA
PQRHSLSEITYLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYA
PQAISKVQPSSYAPQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCM
LGGLSLQEVTSLAMEESQEAKSLHQPLGICTDRSDPNVLHSGEEGTPQYLKGQLPLLSSVQI
EGHPMSLPLQPPSGPCSPSDQGPSPWGLLESLVCPKDEAKSPAPETSDLEQPTELDSLFRGLA
LTVQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

FIGURE 166

MAAAPGLLFWLFWLGALWWVPGQSDLSHGRRFSDLKVC GDEEC SMLMYRGKALEDFTGPDCRF
VNFKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPADETDFVCFE
GGRDDFNSYNVEELLGSLELEDSPPEESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRADS
EDGECAFSESTEGLQGQPSAQESHPTHSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGN
SSPASVEREKT DAYKVLKTEMSQRGSGQCIVIHYSKGFRWHQNLSLFYKDCE

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCAAGAGGCCGGGAAGAGAAGCAAAGC
 GCAACGGTGTGGTCCAAGCCGGGTTCTGCTTCGCCTCTAGGACATACACGGGACCCCCTAA
 CTTCAGTCCCCAAACCGCGACCCTCGAAGTCTTGAACTCCAGCCCCGACATCCACGCGCG
 CACAGGGCGGGCAGGCCAGGTCCGGCGAAGGCAGTCGCGCAGGGGGTGGCAGCTGG
 GCTCGGGCGGCGGGAGTAGGGCCGGCAGGGAGGCAGGGAGGCTGCATATTCAAGAGTCGCGGG
 CTGCGCCCTGGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCA**A**
TGAGCCCGTGGTCTCGCTGCTGGCGCCGCTGCTCTGCCACGGAGCCTCTGCC
 GCCCGTGGTCAGCGCCAAAAGGTGTGTTGCTGACTCAAGCATCCCTGCTACAAAATGG
 CCTACTTCCATGAACTGTCCAGCCGAGTGAGCTTCAGGAGGCACGCCCTGGCTGTGAGAGTG
 AGGGAGGAGTCCTCAGCCTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGC
 AAAACCTGACAAAACCCGGACAGGGATTCTGATGGTGATTCTGGATAGGGCTTGGAGGA
 ATGGAGATGGGAAACATCTGGTGCCTGCCAGATCTTACCAAGTGGTCTGATGGAAGCAATT
 CCCAGTACCGAAACTGGTACACAGATGAACCTCCTGCCAGATCTTACCAAGTGGTCTGATG
 ATCACCAACCAACTGCCAACCTGGCCTGGGGTCCCTACCTTACCAAGTGGAATGATGACA
 GGTGTAACATGAAGCACAATTATTTGCAAGTATGAACCAGAGATTAATCCAACAGCCCTG
 TAGAAAAGCCTTATCTTACAAATCAACCAGGAGACACCCATCAGAATGTGGTTACTGAAG
 CAGGTATAATTCCAATCTAATTATGGTTATACCAACAATACCCCTGCTTTACTGATAC
 TGGTTGCTTTGGAACCTGTTGCTGATAAAAGTAAAGGAAGAACAAAAACTA
 GTCCAAACCAGTCTACACTGTGGATTCAAAGAGTACCAAGAAAAGTGGCATGGAAGTA**T**
AATAACTCATTGACTTGGTCCAGAATTGTAAATTCTGGATCTGTATAAGGAATGGCATCAG
 ACAATAGCTTGGAAATGGCTTGAAATCACAAGGATCTGCAAGATGAACGTAAAGCTCCCT
 TGAGGCAAATATTAAGTAATTCTTATATGTCTATTATTCATTAAAGAATATGCTGTGCTA
 ATAATGGAGTGAGACATGCTTATTGCTAAAGGATGCACCCAAACTCAAACCTCAAGCAA
 TGAAATGGACAATGCAGATAAAAGTTGTTATCAACACGTCGGAGTATGTGTTAGAAGCAAT
 TCCTTTATTCTTCACCTTCATAAGTTGTTATCTAGTCATGTAAATGTATATTGTATTGA
 AATTTACAGTGTGCAAAAGTATTACCTTGCTAAAGTGTGATAAAATGAACGTGTTCTA
 ATATTATTATTATGGCATCTCATTTCATAACATGCTCTTGTATTAAAGAAACTTATTAC
 TGTTGTCAACTGAATTCACACACACAAATATAGTACCATAGAAAAAGTTGTTCTCGAA
 ATAATTCATCTTCAGCTCTGCTTGGTCAATGTCTAGGAAATCTCTCAGAAATAAGA
 AGCTATTCTTAAGTGTGATATAAACCTCCTCAAACATTACTAGAGGCAAGGATTGTCT
 AATTTCAATTGTGCAAGACATGTGCCTTATAATTATTAGCTAAAATTAAACAGATTG
 TAATAATGTAACTTGTTAATAGGTGCATAAACACTAATGCAGTCATTTGAACAAAAGAAGT
 GACATACACAATATAATCATATGTCTCACACGTTGCCTATATAATTGAGAAGCAGCTCTG
 AGGGTTCTGAAATCAATGTGGTCCCTCTTGCCTAAACAAAGATGGTTGTTGGGTT
 GGGATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTCCCTAGCTGTATTAGC
 CTCTGACTATTAGTATAAAAGAGGTCAATGTGGTTGAGACCAGGTGAATAGTCACTATCAG
 TGTGGAGACAAGCACAGCACAGACATTAGGAAGGAAAGGAACTACGAAATCGTGTGAAA
 ATGGGTGGAACCCATCAGTGATCGCATATTGATGAGGGTTGAGATAGAAATG
 GTGGCTCCTTCTGTCTTACCTAGTTCAATGCTTACGCCCTGTTCTCAAGAGA
 AAGTTGTAACCTCTGGTCTCATATGTCCCTGTGCTCCTTTAACCAAAATAAGAGTTCTTG
 TTTCTGGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGFCCR VVSGQKVCFADFKHPCYK MAYFHELSSRVSFQE ARLACES
EGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWS DGSN
SQYRNWYTDEPSCGSEKCVV MYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAP
VEKPYLTNQPGDTHQN VVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKS KGRTKT
SPNQSTLWISKSTRKESGM EV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217